PB-216 230

OPERATIONAL DEMONSTRATION OF OXIDATION DITCH METHOD OF WASTE TREATMENT

Harry M. Kaneshige

Ohio University
Athens, Ohio

21 May 1968

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

PB 216 230

PROGRESS REPORT

Grant No. WPD 100-02-67

Operational Demonstration of Oxidation Ditch Method of Waste Treatment

May 1968

Ohio University ATHENS, OHIO
RESEARCH INSTITUTE

Reproduced by
NATIONAL TECHNICAL INFORMATION SERVICE
U.S. Department of Commerce
Springfield VA 22151

Progress Report

Grant No. WPD 100-02-67

Operational Demonstration of Oxidation Ditch Method of Waste Treatment

Ohio University
Athens, Ohio 45701

Harry M. Kaneshige, Project Director
Associate Professor of Civil Engineering

Period of Report: 1 May 1967 to 30 April 1968
Report Prepared 21 May 1968

Submitted to Federal Water Pollution Control Administration

TABLE OF CONTENTS

		Page
A.	Summary	1
B.	Statement of Progress	
	1. Activities and Accomplishments	2
	2. Project Staff	4
	Appendix	

Summary

A statistical optimization method has been developed for the systems analysis of this project and applied to the data collected during the first seventeen months. The results of the analysis are presented as three separate papers in the Appendix.

Chlorination studies on batch samples in the laboratory indicate that the final effluent from the Somerset plant will require a dosage of about 8 mg/l to yield a residual of 2 mg/l of chlorine. Residuals of 0.5 to 1.0 mg/l produced 99.8% removal of coliform organisms.

Biochemical oxygen demand reaction rate constants of 0.131 per day for raw wastewater and 0.078 per day for final effluent were obtained.

Samples of raw wastewater, final effluent, mixed liquor, return sludge, and stream water were routinely analyzed two or three times weekly. Averages and ranges of values for the various laboratory analyses are presented.

Statement of Progress

1. Activities and Accomplishments

a. Systems analysis

The first year of this project was devoted to preliminary model definition and to identification of needed analytical tools. The oxidation ditch process was divided into four separate but interdependent systems--mechanical, chemical, bacterial, and external environmental.

During the second year, the mathematical techniques for statistical optimization have been refined and applied to data collected from June 1966 to October 1967. Multiple-parameter linear and non-linear regression models and an optimalization model form a basis for the trade-off studies which are the basis for design chart definition. The design charts are to be the ultimate product of this systems analysis.

The results of work in this area are again included as an appendix to this report. The studies are reported in the form of three papers prepared for eventual publication: (1) "Statistical Optimization", (2) "Trade-Off Portraits and Design Charts", and (3) "Polynomial Regression Analysis of Input-Output Models". Software and programming for the IBM 360 computer have been assembled as a separate enclosure.

b. Chlorination

In preparation for chlorination studies, a chlorine contact basin was constructed at the Somerset plant and an "Advance" chlorinator was purchased.

Batch tests were performed in the laboratory on final effluent samples to determine the relationship between chlorine dosage and residual. Calcium hypochlorite and a 15-minute contact period were used. Residual chlorine was determined in accordance with "Standard Methods".[1] The starch-iodine endpoint was used with sodium thiosulfate as the reducing agent. Results of tests on five samples are plotted in Figure 1.

Several of the batch tests for chlorine residual were also analyzed to determine the residual necessary for coliform destruction. After the contact period, this portion of the sample was dechlorinated and tested for the presence of coliforms by the fermentation tube technique. A 99.8% destruction of coliforms occurred when the residual chlorine concentration was between 0.5 and 1.0 mg/l. Assuming that 99.9% removal of coliforms constitutes satisfactory disinfection and that a residual of 2 mg/l is needed to attain this, a dosage of about 8 mg/l will probably be required for Somerset plant effluent.

[1] Standard Methods for the Examination of Water and Wastewater, 12th ed, published jointly by American Public Health Association, American Water Works Association, and Water Pollution Control Federation, 1965.

Studies have recently been resumed for verification of last summer's results.

c. B.O.D. reaction rate constants

The biochemical oxygen demand (B.O.D.) reaction rate constants of the raw wastewater and final effluent were determined with a Weston and Stack dissolved oxygen probe. The values of the reaction rate constants were calculated by Thomas' method.[2] Five runs of approximately 20-day durations were made on each type of sample. For the raw wastewater, the constant ranged from 0.065 to 0.205 per day, with an average of 0.131 per day. For the final effluent, the constant ranged from 0.028 to 0.157 per day, with an average of 0.078 per day. All of these reaction rate constants are for base 10.

d. Laboratory analyses

Samples of raw wastewater, final effluent, mixed liquor, return sludge, stream just above the plant, and stream just below the plant were collected and analyzed twice or thrice weekly. The results are shown in Table 1. These results must be accepted as preliminary, because the punching of the data onto cards has not been completely checked as yet.

During the period from June 1966 to October 1967, the average B.O.D. loading based on the average flow of 100,760 gallons per day and average B.O.D. concentration of 267 mg/l in the raw wastewater was 224 lb/day or 8.2 lb/day per 1000 cu ft of ditch volume.

e. Automatic analyzer

The Civil Engineering Department has ordered a Technicon Auto-Analyzer which will be used on this project during the coming year. This instrument is expected to permit the analysis of a large number of samples, so that replication and studies of short term fluctuations of several chemical constituents will be facilitated.

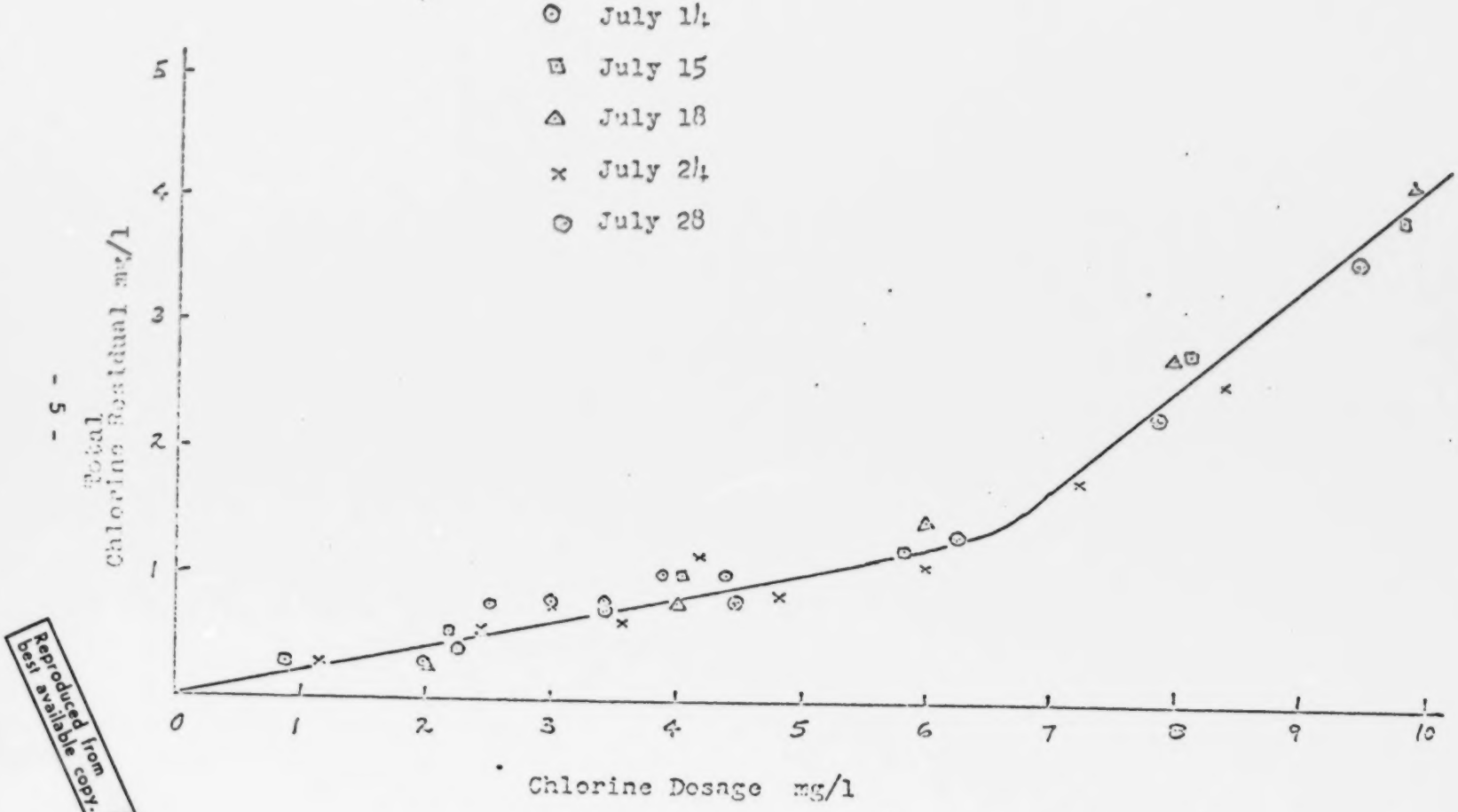
[2] Thomas, H. A., Jr., "Analysis of the Biochemical Oxygen Demand Curve", Journal of the Water Pollution Control Federation, Vol 12, p 504, 1940.

2. Project Staff

Harry M. Kaneshige, Associate Professor of Civil Engineering, Project Director, devoted approximately 36% of his time to this project from June 1967 through April 1968.

Edward H. Gamble, Professor of Industrial and Systems Engineering, devoted approximately 28% of his time to this project from June 1967 through April 1968.

Reuben M. Olson, Professor of Civil Engineering, devoted less than 1% of his time to this project during this past year.



Reproduced from best available copy.

Figure 1

Chlorine Residual versus Chlorine Dosage

Table 1
Summary of Laboratory Analyses
June 1966 - October 1967

		Raw	Final	Stream Above	Stream Below
Alkalinity mg/l	Average	268	66	112	81
	Range	122-376	13-157	43-172	32-126
B.O.D. (unfiltered) mg/l	Average	267	8	2	4
	Range	66-684	2-76	0-21	0-24
B.O.D. (filtered) mg/l	Average	98	3	--	--
	Range	12-185	1-11		
C.O.D. mg/l	Average	583	43	--	--
	Range	129-972	3-473		
Ammonia mg N/l	Average	34	0.30	0.07	0.11
	Range	6-75	0-8	0-0.53	0-1.3
Nitrite mg N/l	Average	0.05	0.06	0.01	0.04
	Range	0-1.06	0-0.61	0-0.15	0-0.36
Nitrate mg N/l	Average	0.4	26	0.9	12
	Range	0-7.6	4-54	0-3.6	2.6-39
pH	Average	7.2	6.9	7.5	7.2
	Range	6.4-7.6	5.8-7.8	5.8-8.6	5.8-8.2
Total Phosphate mg PO_4/l	Average	53	40	0.7	17
	Range	11-95	5-76	0-5.2	1.2-53
Suspended Solids mg/l	Average	348	22	13	21
	Range	114-1254	0-532	0-343	0-368
Total Solids mg/l	Average	911	584	299	411
	Range	494-1809	433-1095	197-655	211-748

		Mixed Liquor	Return Sludge
Suspended Solids mg/l	Average	6860	9320
	Range	2400-10060	207-25890
Suspended Solids % Volatile	Average	61	61
	Range	48-88	49-83
Percent Volume (30 min. settling)	Average	72.7	
	Range	19-97	

APPENDIX

The Appendix of this report has been prepared by Dr. E. H. Gamble, Department of Industrial and Systems Engineering

TABLE of CONTENTS for the SYSTEMS ANALYSIS SECTION

page

I. Introduction..... 1

A. Background 1
B. Format of the Analytical Studies 4

1. Study A 4
2. Study B 5
3. Study C 7

II. Summary of Results 9

III. Forecast for the Third Year of the Program ..11

IV. Paper : STATISTICAL OPTIMIZATION ..... 40 pages

V. Paper : TRADE-OFF PORTRAITS and DESIGN CHARTS .36 pages

VI. Paper : POLYNOMIAL REGRESSION ANALYSIS of INPUT-OUTPUT MODELS ..... 5 pages

ENCLOSURES :

IBM :

1. Software programming and catalogue of measurement data from laboratory analysis of plant data.
2. Software programming and analysis for multiple parameter linear and curvilinear regression modelling. Includes a new Z_i^2 confidence test.
3. Softare programming for multiple around - the- season system adaptivity analysis. Eight (8) seasonal operating sets of conditions considered. Trade-off studies made and portraits generated. These form family plots of the eight J conditions.
4. Sample regression modeling utilizing the multi-term polynomial of highest for practical and satisfactory correlation.

Second Annual Report

SYSTEMS ANALYSIS SECTION

I. Introduction

A. Background

The first year of our demonstration program was devoted to preliminary model definition and identification of needed analytical tools. Our studies indicate the desirability of assigning the total system to four participating and interdependent systems:

1. A bacterial system
2. A chemical system
3. A mechanical-physical system
4. An environmental system

A preliminary model was defined for each of these systems including those parametric-variables which are susceptible to experimental measurement either directly or indirectly. The measurements of the bacterial-chemical variables are found from selected "raw" and "final" system fluid samples. Laboratory analyses of these samples yield considerable discrete-value data compatible with the total system season and diurnal time of collection. These laboratory data are representative of the indirect method of data collection.

Direct measurement is made of the environmental variables such as air and fluid temperatures, earth temperatures, atmospheric pressure, etc. The rotor angular velocity and transport velocities near the rotors are found indirectly from the measurement data for the electric power consumption

rate of the electric motors driving the rotors and paddles. Of course, the dependency relationship of rotor velocities upon electric power rate must have been established previous to our detailed system data analyses.

It will be our planned program to collect pertinent data on a selected basis. If a year is divided into six (6) week periods, the eight (8) periods for one total year provides us with a statistical data level of finite duration in system time and a large enough "lot" and sample space to yield good prediction of system population/ensemble characteristics.

Our attempts to utilize multiple parametric linear and non-linear regression methods to analyze the sample data for each level or time period will show the need for some technique extensions. Mathematical model definition for systems of n-space as large as 27 parameters requires the analyst to utilize a large capacity digital computer and provide careful software programming. Care is mandatory if the intermediate data is in a form to facilitate decision-making based upon engineering judgment assisted by Markov process techniques.

The regression models form a first step towards practical design model definition for our complex system. An optimalization model analysis should establish the degree of accuracy of the regression parametric mean values as system population central values. These two models form a basis for the design trade-off studies which are the basis for design chart defin-

ition. The design charts are the hoped-for end-product of our system analyses program. Design charts to be practical are given in a form permitting selection of the physical configuration parametrics and capacity potential values for a range of system sizes and applications. Considerable knowledge about the complex system should be derivable from a critical analytical study of these model and design charts -- in very practical engineering terms.

Our second year of this project reports upon the collections and analyses of eighteen (18) months of experimental sampled data. Extensive laboratory data were assembled and cataloged. We will report upon the results of our application of several regression techniques, their needed extensions and the resultant definition of trade-off plots. These trade-off parametric portraits will give an empirically-derived objective function as the selection policy generator for design-operating values.

As we draw to a close of our second year of the project, we may derive a practical level of confidence in our capabilities to analyze a most complex waste-water processing system, based entirely upon experimentally collected data on an existing system. Inadequate design data and undefined parametric-selection policies were available upon which to base any of our demonstration analysis. Only an existing and in-operation plant of unknown characteristics and capabilities was available. Yet this unknown operating plant was found to

yield a wealth of empirical data upon which to base our studies. Moreover, as an operating system, it has continuously provided a good level of confidence to each observer in the basic concepts of its design.

B. Format of the Analytical Studies

To bring realistic order to the measurement, laboratory analyses and cataloguing of the bacterial-chemical parametric data, a computer software programming approach was developed. The results of this catalogue are seen in the accompanying enclosure. A detail description and guide to the catalogue is given in the preface and introduction to the document.

Two (2) analytical studies are included in the report. These may be described as:

1. "Statistical Optimization". This paper describes a processing technique utilizing experimental sample data, forming linear and non-linear regression models for the dependency of a major system output variable upon a large but representative number of bacterial-chemical parameters. The regression models are improved through use of a reverse Markov chain to find corrected or near-optimal expected values differing from those defined as the sample arithmetic means. These expected values satisfy the multiple regression models found via a method of least squares. A testing measure is given for determining the

need for including additional non-linear terms to each model iteration.

As justification for limiting our regression models to second order terms, in addition to the linear terms, a separate study was conducted for function matching of several of the more significant parametrics and their dependencies. A graphical matching (eye-ball engineering) is compared to the empirical function matching via computer methods. The sensitivity coefficients for the third power and higher terms may be considered as negligible in comparison with the coefficients of the linear and squared terms of the parametrics.

For each major set of environmental conditions, a compatible, consistent set of expected values and their equivalent set of optimal values may be found from a combined regression and optimization technique. This technique is based upon the selected criteria or objective functions and companion selection policies. The objective functions are determined by empirical techniques including graphical and computation methods.

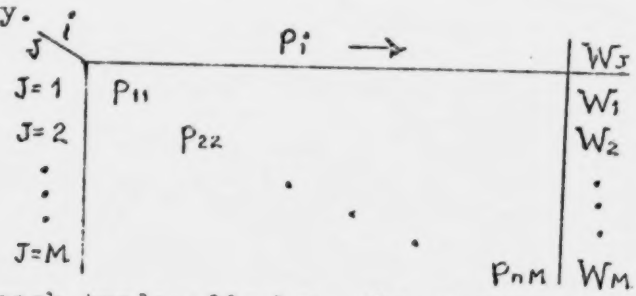
2. "Trade-off portraits and design charts" are determined analytically through use of the Theory of Inequalities and Convex Functions. A trade-off portrait for the output variable (sewage flow-rate)

against each system independent variable may be plotted for each environmental condition set. A first set of parametric values for each condition or $\{p_i\}_J$ are, in fact, the design operating set for the M different J condition sets. A curve passed through the M points is identifiable as the two-dimensional projection of the n parametric objective function.

Our study concerns the a) use of the experimentally derived set of parametric values ($j = 1, 2, \cdots, m$) for each of the $M - J$ environmental conditions, b) the definition of the regression model, c) the study of the M models as the constraint definition of an optimal model with objection criteria $\Omega(\chi_i)$, a derived set of optimal values $\{p_i\}^\circ$. This is a optimal set of parametric values

$$\{p_i\}^\circ = (p_{iJ})^\circ_{J=i}$$

It forms the diagonal values of a parametric matrix array.



For each trade-off plot, the y versus χ_R for the several J conditions forms system condition functions $y_J(\chi_R, \bar{\chi}_i)_{i \neq R} = \Phi_J(\chi_R, \chi_i)_{i \neq R}$. The set $\{\chi_{RJ}\}$ form the Rth column of the para-

metric matrix. There will be M such design-operating values. The locus of this value set forms the objective criterion $\Omega(x_R, \bar{x}_i)$ where $(\bar{x}_{iJ})$ are fixed at their central or expected values for each J criterion and the $\{x_{RJ}\}^\circ$ are restrictions upon the permissable values of x_R. The $\Phi_J(x_R, x_i)_{i \neq R}$ functions are identical to the curvilinear contributions $\left[f_R^{(2)}(x_R)\right]_J$ found for the multiple regression modeling, where $\{x_i\} = \{x_{iJ}\} = \{\bar{x}_i\}_J$, and x_R is variable.

The two dimension plot of the loci of these values is not the $\Omega(x_R, \bar{x}_i)_{i \neq R}$ function but is the projection of this function upon the y vs. x_R two dimensional surface. $\Omega(x_R, \bar{x}_i)_{i \neq R}$ is a curve in the n-space of x_i constrained to the $\{\bar{x}_i\}_J$ intersection of $\Omega(x_{RJ})$ with each $(x_i)_J$ spacial location. Hence, $\Omega(x_R, \bar{x}_{iJ})$ is not constrained only to the $\{x_{RJ}\}^\circ$ set but may vary over any of the x_R range of values permissible with physical realizability and $x_R \geq 0$. The $\{x_{RJ}\}^\circ$ form a sub-set of χ_R. The $\{x_{RJ}\}^\circ$ must be compatible with the $\{x_{iJ}\}$ selected sub-sets.

A second trade-off plot will show y vs. x_q and the $y_J(x_q, \bar{x}_i)_{i \neq q} = \Phi_J(x_q, \bar{x}_i)_{i \neq q}$ form a second set of system condition functions. Again, a cut of $\Omega(x_i)$ corresponding to $\Omega(x_q, \bar{x}_{iJ})$ has a two dimensional projection

on the y vs x_q plane. The locus of $\{x_{qJ}\}^{\circ}$ and $\{x_{iJ}\}^{\circ}$ and the x_q values satisfying $\Omega\binom{x_i}{x_{iJ}}$ will give values for $\Omega(x_i)$ in terms of the y variable.

The reader is referred to pgs. 8 through 19 of Papers A and B, for a set of graphical models for the trade-off plots and relationships. Our discussion is brief in this section, since the purpose is to briefly state the background for our program needed to supplement the work reported for the second project year. Any mathematical derivations needed for Paper A or Paper B will be included in the body of material of the appropriate paper.

II. Summary of Results (at time of Second Annual Report)

A wealth of experimental data can and has been assembled. Because of the system complexity, however, sample values for many significant parameters can not be found directly. It would be virtually impossible to assemble sample data for all parameters for every hour and every changed environmental condition. Hence, our analyses are approximate and the degree of correlation adequate but not optimum.

The proposed technique for statistical optimization has been found to have several practical restrictions. These restrictions are associated with or directly attributable to the needed form for software programming. Our present day digital computers can handle complex problem solutions such as multiple regression (linear and/or non-linear) programs only in a rather restricted way. The variables are introduced into an iteration step inversely (firm or not-so-firm) to the values of the linear model single correlation coefficients and some judgment test such as the f or t hypothesis method. The major or final last-word test still seems to be engineering judgment or common sense valuation of the parametric values and relationships in their physical meaning. With appropriate and frequent exploration of engineering judgment, the statistical optimization technique appears to have value. Primarily, it is seen as a direct step in the formulation of trade-off portraits. Moreover, the improvement and/or rejection of non-linear modeling terms for multiple regression

correlation is straight-forward. A testing hypothesis closely related to the confidence level test of the frequency distribution of sample data, (the use of the standardized normal distribution through the blessing of the Central Limit Theorem), shows merit.

The trade-off parametric portraits can be based upon experimentally derived sample data, statistical regression model techniques, and the analytical formulations of the Theory of Inequalities and Convex Functions. The addition or subtraction of the contribution to y (the dependent variable) by any single or combination of independent parameters can be performed easily. The results are directly interpretable in terms of system physical definition. The empirically found selection criterion curve is consistent with the analytically-derivable objective criteria and companion selection policies. There is no need for an iterative process of juggling a set of portraits to find a consistent set. The analytical methods described in Paper B are sound and not too difficult to apply. Even a guided set of semi-complex steps in an analytical technique is preferrable to a hit-or-miss graphical process of iterations.

The most valuable aspect of the trade-off techniques presented is the directness for finding and identifying Design Charts as required for new system design, consistent with its application and environmental conditions.

III. Forecast of the Third Year Program

Our third year program will concentrate upon direct definition of consistent sets of system design charts and required installation specifications.

The results of our second year efforts were in development of regression model techniques and the software programming to exploit the techniques. The techniques were used for models both large and small in space size. Two-dimensional models for only two parameter sub-set must be defined. The direct influence of the environmental sub-set of variables upon each of these models will be quantified in order to establish the degree of system adaptivity.

A number of optimalization studies (primarily analytical) will be used to determine both the "System Effectiveness" and its companion the "Cost Effectiveness." By the first term we mean the predicted (via statistics and probability) absolute magnitude of output variable variations and the time rate of deterioration and degradation. To find these quantified values, the designer-evaluator must carefully define his performance criteria, pattern of tolerance assignments, and their propagation. Essentially, this entails the definition of realizable parametric tolerance profiles. These performance studies must of necessity be related to reliability seen as a value standard under strict location and operating condition constraints. The design equipment models have been derived in terms of system functional modeling rather than the more

common availability-mission readiness probabilistic methods. Our method is more advantageous, since the sub-system and system variable variation may be derived from parametric value variations on a statistical basis. From the experimental data, we have found directly or indirectly the sample variances for each parameter. These are available for our effectiveness analyses.

"Cost Effectiveness" is comparison with "System Effectiveness" involves more than performance and adherence to technical design requirements. The practicality of the engineering decisions of configuration, duty cycle for input utilities or external resources, adaptivity, accessibility and maintainability potentials serve to define relative system and sub-system complexity. As an example, where we consider power consumption level as one measure of complexity, we should consider the duty cycle or length of operating time to find the total electric power required.

One approach to "C.E." places the integration aspects solely in the hands of the sponsoring system customer. The contractor is greatly restricted in exploiting trade-off techniques to obtain a better system design. An integrated design study is often hampered, if not discouraging. For a large system such as we are analyzing, one finds the nature of the data required seems overwhelming at first glance. However, as we shall show, many common factors are recognizable. Operational Effectiveness is one of the more important factors. One can transfer the system detail design and system integration to the builder-

contractor provided:

a) He can be furnished with a set of integrated figures of merit -- the desirable selection criteria needed for measuring the degree of appropriateness of system specified equipment.

b) A program for proper utilization of specified system equipment (where the selection criteria and design charts can be formulated).

Broader latitude of control within the scope of a system operating in a complex environment must be identified and permitted. In addition to the "S.E." features and factors, the maintenance resources investment may be determined. Such studies for the prediction of relative operational and maintenance costs are scheduled for our third year program. The detail planning for these has not been undertaken as yet.

Study A

"Statistical Optimization"

(Applied to Demonstration and Analysis of the Somerset Oxidation-Ditch Method for Waste-Water Processing)

by E. H. Gamble, Ph.D.
Professor of Systems Engineering
Ohio University
Athens, Ohio

"STATISTICAL OPTIMIZATION"

Introduction

When our engineering situation is one of evaluating a complex experimental system, specialized techniques must be adopted in order to characterize and model the situation and compatible system. An oxidation-bacterial waste water processing plant provides an excellent application for the utilization of sampling data regression techniques for modeling when made compatible with the more common system qualitative-quantitative characterization methodology.

A complex system of many variables and several inter-dependent sub-systems presents difficulties in analysis, even when all parameters and sensitivities can be measured. This is particularly true where the system is operated under many controllable sets of static-dynamic conditions. The addition of limited measureability and limited control of operating conditions increases the demands upon our methodology and techniques -- and our engineering judgment. A statistical approach appears logical, but some deficiencies in the methodology are readily found when applied to complex process-plants. It will be the objective of this paper to report one study and its results to help correct some deficiencies through clarification of the terminology and methodology. The system to be evaluated permits only experimental sampling of specific parameters

at numerous seasonal and diurnal environmental conditions. When taken over too broad a spectrum of seasonal environmental conditions, it is difficult to find parametric sensitivities and system influence-effect weighting coefficients. Hence, one single model appears impractical. Multiple application of some form statistical optimization to define each major system operating conditions helps overcome this difficulty.

Utilization of Multiple Parameter Regression Technique

In statistical form, a system or sub-system constraint relationship is definable as:

$$y_J \leq \beta_{oJ} + f_{1J}(x_1) + f_{2J}(x_2) + \cdots + f_{NJ}(x_N) \qquad (1.01)$$

where $f_{iJ}(x_i)$ is a regression showing

a) Linear dependence of y on x_i for each J condition set

$$f_{iJ}(x_i) = \beta_{iJ}\, x_i \qquad (1.02)$$

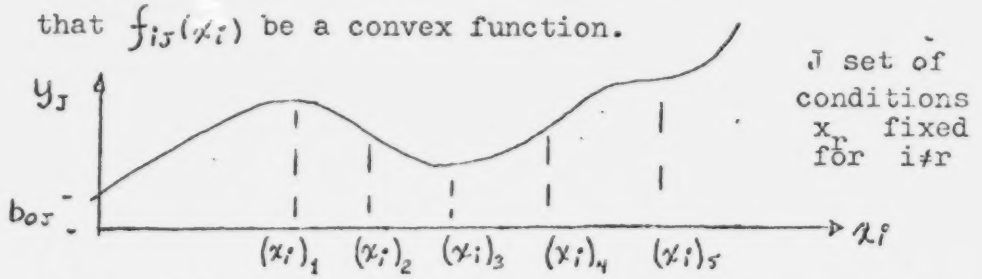
b) Linear independence (or a constant contribution)

$$f_{iJ}(x_i) = \beta_{iJ}\, \bar{x}_i \qquad (1.03)$$

c) Curvilinear

$$f_{iJ}(x_i) \approx \beta_{iJ} \cdot x_i + \beta_{iJ}^{(2)} \cdot (x_i^2) + \cdots + \beta_{iJ}^{(n)} (x_i^n) \qquad (1.04)$$

but for practical systems even a cubic regressional dependence must be taken in regions of x_i to assure that $f_{iJ}(x_i)$ be a convex function.



(c-1) for $0 \leq x_i \leq (x_i)_1$, $f_{iJ}(x_i)$ is essentially parabolic and limited to $b_{oJ} \leq f_{iJ}(x_i) \leq f_{iJ}(x_{i,1})$

(c-2) for $(x_i)_1 < x_i \leq (x_i)_2$

$$f_{ij}(x_i)_1 > f_{ij}(x_i) \geq f_{ij}(x_i)_2$$

(c-3) for $(x_i)_2 < x_i \leq (x_i)_3$
parabolic of negative sensitivity $-\left(\frac{\partial y}{\partial x_i}\right)_j$

(c-4) for $(x_i)_3 < x_i \leq (x_i)_4$
parabolic of positive sensitivity $\left(\frac{\partial y}{\partial x_i}\right)_j$

(c-5) for $(x_i)_4 < x_i \leq (x_i)_5$
parabolic of negative sensitivity $-\left(\frac{\partial y}{\partial x_i}\right)_j$

(c-6) for $(x_i)_5 < x_i \leq (x_i)_N$
parabolic of positive sensitivity $\left(\frac{\partial y}{\partial x_i}\right)_j$

These regions and their boundaries are necessary to avoid multiple-valuedness and to assure convexity (which is shown to be so necessary for consistent-compatible optimization in systems design and analyses).

The interdependency or constraint relationship (1.01) is an equality only when all M system parameters are taken into account and the $\{f_{ij}(x_i)\}$ set are analytically defined. For the selected set of N x_i parameters, where $N < M$

$$(y)_j = \beta_{oj} + \sum_{k=1}^{N} \beta_{kj} x_{kj} + \sum_{k=N+1}^{M} \beta_{jk} x_{jk} + \text{higher order terms} \qquad (1.05)$$

For one set of system operating and environmental conditions equation 1.05 is our regression model. If only N x_i parameters are included:

$$y_j = \beta_{oj} + \sum_{k=N+1}^{M} \beta_{jk} \bar{x}_{kj} + \sum_{i=1}^{N} f_{ij}(x_i)$$

$$= B_{oj} + \sum_{i=1}^{N} f_{ij}(x_i)$$

where $\beta_{kj} \cdot \bar{x}_{kj}$ is found from a single linear regression analysis for each of the $M - (N+1)$ variables. Hence (1.06) and (1.07) are each an approximation to the model derivable from a multiple parameter curvilinear analysis. For most practical physical

systems $(B_o)_J$ will be a threshold value differing from zero in value. Moreover, rarely will the systems engineer have access, measurement equipment, or time to take experimental data for all of the $(M+1)$ parameters including y. He will be forced to select a limited set of N x_i which appear to have system significance and influence.

A system parametric study is interested in establishing the functional dependence of an output variable y upon each of the $\{x_i\}$ parameters when the other $(N-1)$ variables are held at their design operating values. A portrait of functional curves (one for each J set of conditions) becomes a trade-off study plot for optimal solution of the set of parameter design-values for each J set. There will be N such portraits for the N - x_i and the one y variables. Since we are for this situation beginning with experimental data, measured for the $N+1$ variables for m different j datapoints, we must determine what relationship the expected regression values have to our desired set of design operating values. A multiple regression analysis will determine the model.

$$(y)_J = (B_o)_J + \sum_{i=1}^{N} f_{i,J}(x_i) \qquad (1.08)$$

and

$$(\bar{y})_J = (\bar{B}_o)_J + \sum_{i=1}^{N} f_{i,J}(\bar{x}_i) \qquad (1.09)$$

where the arithmetic statistical means are:

$$\bar{x}_i = \frac{1}{m} \sum_{j=1}^{m} (x_i)_j \quad , \quad \bar{y} = \sum_{1=1}^{m} (y)_j \qquad (1.10)$$

Knowing (1.09) and (1.10) results where $m > 30$, the $\bar{x}_i$ and $\bar{y}$ are unbiased estimates of the expectations. These values provide a platform for empirical determination of the $\{f_{i,J}(\bar{x}_i)\}$ and $\{f_{i,J}(x_i)\}$ sets.

-4-

For a particular multiple regression analysis of a system of $(N+1) = 26$ variables where $j = 1, 2, \cdots, 44$ as experimental sets of data samples over a large population of values and conditions, we determined ($i = 24$ is the last variable in the regression iteration)

$$(\bar{y})_{\substack{25th \\ iteration}} = (B_0)_{\substack{constant \\ 25th \\ iteration}} + \sum_{\substack{i=1 \\ i \neq 24}}^{25} f_i(x_i) + f_{24}(x_{24}) \qquad (1.11)$$

If we set:

$$(\bar{y})_{\substack{25th \\ iteration}} = \frac{1}{44} \sum_{j=1}^{44} (y)_j \qquad (1.12)$$

and $$f_{\substack{i \neq 24}}(x_i) = (\beta_i)_{\substack{25th \\ iteration}} \cdot x_i \qquad (1.13)$$

$$, \quad \bar{x}_i = \frac{1}{44} \sum_{j=1}^{44} (x_i)_j \qquad (1.14)$$

then $$f_{24}(x_{24}) = \bar{y} - B_0 - \sum_{\substack{i=1 \\ i \neq 24}}^{25} (\beta_i)_{25th} \cdot (\bar{x}_i) \qquad (1.15)$$

Utilizing the value of $f_{24}(x_{24})$, in turn we shall seek a second estimate for $f_{\substack{i \neq 24}}(x_i)$ for each i.

Or:

$$(\bar{y})_{\substack{24th \\ iteration}} = \left[(B_0)_{\substack{constant \\ 24th \\ iteration}} + f_{24}(x_{24})_{\substack{from \\ 1.15}} \right] + \sum_{i=1}^{25} (\beta_i)_{\substack{24th \\ iteration}} \cdot \bar{x}_i$$

$$+ f_{23}(x_{23}) \qquad (1.16)$$

where $i = 23$ is the last variable introduced into the 24th iteration and $\bar{x}_i$ is found from (1.14).

The Reverse Markov Process is essentially a technique for finding the $f_i(x_i)$ determined by application of a multiple parameter regression analysis to experimentally derived data for a complex system. It appears somewhat of a bootstrapping process

since we start with the most recent data of the last iteration of the regression process. We select the $f_i(x_i)$ functional contribution of the last variable introduced by the iteration steps of the regression. This contribution $f_i(x_i)$ for our example is for the $i = 24th$ variable and we compute it from (1.15) . Inspection of (1.15) shows that for all variables other than the 24th we substituted the arithmetic mean values taken from the entire population study and the value B_o and B_i's are the constant term and coefficients of the last iteration (for our example -- the 25th step). With our computed value for $f_{24}(x_{24})$ we proceed to the second step of our Reverse Markov chain and utilize (1.16). The arithmetic means for all i's except 24 and the variable introduced into the next to last step of the regression $(i = 23)$ are used. For this link the constant term and coefficients of the next to last regression interation are employed together with $f_{24}(x_{24})$ from (1.15).

With our Markov estimate for $f_{24}(x_{24})$ and the computation of $f_{23}(x_{23})$ from (1.16), we proceed backwards up the iteration steps to find $f_{22}(x_{22})$ and each $f_i(x_i)$. In the 21st iteration step variable x_9 was introduced.

For example:

$$(\bar{y})_{\substack{23\,RD \\ \text{iteration}}} = \Big[B_{o\,\substack{\text{constant} \\ 23\text{rd} \\ \text{iteration}}} + f_{24}(x_{24})_{\substack{\text{from} \\ (1.15)}} + f_{23}(x_{23})_{\substack{\text{from} \\ (1.16)}} + f_{22}(x_{22})_{\substack{\text{from} \\ 1.16 - \text{step } 22}} \Big]$$

$$+ \sum_{i=1}^{25} (\beta_i)_{23RD\ \text{iteration}} \cdot \bar{x}_i + f_9(x_9) \qquad (1.17)$$

We proceed to find the 24 terms of contributions:

WE PROCEED TO FIND THE 24 TERMS OF CONTRIBUTIONS : $\{f_i^{(1)}(x_i)\}$ in the reverse of their order in the multiple regression computation. For our study the parabolic contribution per initial variable should be added to the linear in order to find the total curvilinear contribution.

Tables I-1, I-2, and I-3 give these values for three bounding limitations upon the total contributions. α = 1.0, 0.1, and 0.01. Where $(y-b_o)_J = \alpha \sum_{i=1}^{16} f_i^{(2)}(x_i)$

(1.18)

	x_i	$(x_i)^2 = x_q$	$f_i^{(1)}(x_i)$	$f_q^1(x_q)$	$f_i^{(2)}(x_i) = f_i^{(1)}(x_i) + f_q^{(1)}(x_q)$
i =	1	18 = q	+ 14586.56	- 2779.59	+ 11806.97
	2	19	- 41868.55	+64192.49	+ 22323.94
	6	20	-105998.94	- 1035.81	-107034.75
	12	21	-694207.87	+662669.87	- 31538.00
	13	22	+ 11202.31	- 825.00	+ 10377.31
	14	23	- 23903.25	+ 860.06	- 23043.19
	15	24	- 140.75	+ 175.69	+ 34.94
	16	25	58265.44	- 54364.75	+ 3900.69

An Hypothesis Test for Correctness of Linear and/or Curvilinear Dependency Determined Via Regression Methods

For multiple parameter regression modelling, one has need for more than the single linear correlation coefficients and a total model correlation coefficient as a testing factor for appropriateness. We will define a testing factor which provides a probability confidence level for the improvement model resulting from selective extensions of our first linear model to a curvilinear one made up of linear and/or parabolic dependencies. Our test factor must be applied to each significant (if not independent) variable.

For a linear multiple parameter regression we would define $$y_i' = b_0 + b_1x_1 + b_2x_2 + \cdots + b_nx_n \tag{1.19}$$ with an anchor point of coordinates $(\bar{y}, \bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n)$ the values are quantified at the arithmetic means. For this (n+1) dimension point set:

$$\bar{y} = b_0 + b_1\bar{x}_1 + \cdots + b_n\bar{x}_n. \tag{1.20}$$

In the previous section of this paper we defined q additional variables equivalent to the square of selected variables x_i, (i=1, 2, ---, n). Moreover:

$$(y_i)^1_{\text{curvilinear}} = B_o + \sum_{i=1}^{n} f_i^{(1)}(x_i) + \sum_{i=n+1}^{n+q} f_i(x_i)$$

$$= B_o + \sum_{i=1}^{n'} f_i^{(1)}(x_i) + \sum_{i=n'+1}^{n'+q} f_i^{(2)}(x_i)$$

where n' is a selected subset of n; $n' \in n$ and (1.21)

$$\sum_{i=n'+1}^{n'+q} f_i^{(2)}(x_i) = \sum_{i=n'+1}^{n} \left[f_i^{(1)}(x_i) \right] + \sum_{i=n+1}^{n+q} \left[f_i^{(2)}(x_i) \right]$$

(1.22)

Our proposed test is somewhat similar to the z value of the equivalent standard normal distribution via the Central Limit Theorem applied to a sample when the population characteristics are assumed. (See Bartee-Reference A.) Our suggested test factor is :

$$z_i^* = \frac{1}{(b_i)_{\substack{\text{LAST} \\ \text{ITERATION}}} \cdot \substack{\text{MAGNITUDE} \\ \text{JUDGMENT} \\ \text{FACTOR}}} \cdot \frac{(b_i)_{\substack{\text{LAST} \\ \text{ITERATION}}} \overline{x_i} - f_i^{(2)}(x_i)_{\substack{\text{LINEAR \&} \\ \text{CURVILINEAR}}}}{\sigma_{x_i} / \sqrt{m}}$$

(1.23)

where $\overline{x_i}$ = arithmetic mean

$$= \frac{\sum_{j=1}^{m} x_{i,j}}{m}$$

(1.24)

$$\sigma_{x_i} = \hat{S} = \left\{ \frac{\sum_{j=1}^{m} (x_j - \bar{x})}{(m-1)} \right\}^{1/2}$$

(1.25)

(b_i) last iteration = regression coefficient for x_i variable as found in last iteration of curvilinear regression analysis.

$f_i^{(2)}(x_i)$ = contribution function value for the x_i variable including both linear and parabolic terms found from reverse Markov Process.

A value for Z_i greater than one (1) is needed for a probability confidence level of 84% or more that our present level of development for the system regression model provides a good match of the experimental data sample and would yield excellent prediction.

Two sets of appropriate data are given in the accompanying tables. (The complete regression analysis is given in our folder of computer data and data processing).

i,ℓ	x_i	$\overline{x_i}$	(ρx_i) single correlation	$f_i^{(1)}(x_i)$	$f_i^{(2)}(x_i)$	$\hat{z}_i^*$	$f_i^{(2)}(x_i)/\bar{y}$ % contribution
1-18	X_1	0.12	0.328	14586.56	+11806.56	6.4180	+42.02
2-19	X_2	59.78	-0.452	-41868.55	+22323.94	36.8693	+79.46
6-20	X_6	266.17	-0.371	-105998.94	-107034.75	68.3927	-381.00
12-21	X_{12}	591.54	-0.438	-694207.87	-31538.00	61.7059	-112.26
13-22	X_{13}	16.52	-0.020	+11202.31	+10377.31	16.4718	+36.93
14-23	X_{14}	61.63	0.383	-23903.25	-23043.19	24.3809	-82.02
15-24	X_{15}	0.06	-0.227	-140.75	+34.94	6.4837	+0.12
16-25	X_{16}	27.39	-0.092	+58265.44	+3900.69	16.2681	+13.88

Table I - 1

(Judgment Factor Equals One (1))

$P(\hat{z}_i^*) = 1.000$

i,q	x_i	$\bar{x}_i$	(ρx_i) single correlation	$f_i^{(1)}(x_i)$	$f_i^{(2)}(x_i)$	$\bar{z}^*_i$	$P(\bar{z}^*_i)$	$f_i^{(2)}(x_i)/\bar{y}$ % contribution
1-18	X_1	0.12	+0.328	+1450.88	+959.30	+0.6404	.739	1.09
2-19	X_2	59.78	-0.452	-4190.11	+2226.60	+7.6419	1.000	2.50
6-20	X_3	266.17	-0.371	+7365.00	+7263.60	+0.4822	.684	8.20
12-21	X_4	591.54	-0.438	-68904.81	-3144.70	+6.1700	1.000	3.60
13-22	X_5	16.52	-0.020	+1108.19	+1033.80	+1.6428	.950	1.18
14-23	X_6	61.63	+0.383	-2390.71	-2296.10	+2.4396	.993	2.25
15-24	X_7	0.06	-0.227	-18.81	-6.70	+0.6232	.732	0.008
16-25	X_8	27.39	-0.092	+5826.50	+388.00	+1.6273	.948	0.450

Table I - 2

(Judgment Factor Equals 0.1)

37

i,c	x_i	$\bar{x}_i$	(ρx_i) single correlation	$f_i^{(1)}(x_i)$	$f^{(2)}(x_i)$	$\bar{Z}_i^*$	$P(\bar{Z}_i^*)$	$f_i^{(2)}(x_i)/\bar{y}$ % contribution
1-18	X_1	0.12	+0.328	+160.66	+131.50	+0.0688	0.520	0.150
2-19	X_2	59.78	-0.452	-421.74	+221.00	+0.3546	0.630	0.250
6-20	X_6	266.17	-0.371	+723.86	+714.60	+0.0478	0.520	0.816
12-21	X_{12}	591.54	-0.438	-6958.59	-303.20	+0.6174	0.730	0.350
13-22	X_{13}	16.52	-0.020	+98.39	+90.80	+0.1441	0.550	0.100
14-23	X_{14}	61.63	+0.383	-238.96	-220.40	+0.2414	0.600	0.250
15+24	X_{15}	0.06	-0.227	+15.55	+15.30	+0.1025	0.540	0.017
16+25	X_{16}	27.39	-0.092	+582.42	+37.90	+0.1631	0.560	0.045

Table I - 3

(Judgement Factor equals 0.01)

38

Parabolic Limitation Study

Our curvilinear models have been limited to parametric parabolic dependency. This limitation decision resulted from an independent study--reported as Study C. This study concerned the relative improvement in individual dependency plots of y versus x_i. The reduction in residual error with increasing higher order contribution terms became unprofitable beyond the second power of x_i. Hence we limit

$$f_i^{(2)}(x_i) = f_i^{(1)}(x_i) + f_q^{(1)}(x_q) \qquad (1.26)$$

linear

for $q = (x_i)^2$

$$= (b_i)_1 (x_i) + (b_i)'_2 (x_i)^2 \qquad (1.27)$$

$$f_i^{(1)}(x_i) = (b_i)_1 x_i \qquad (1.28)$$

To illustrate this phenomena and decision justification, the IBM computer makes y vs x_i plots for the several curvilinear models. Numerical results are provided to verify the portrait values and conclusions.

Comments Upon Z_i^* Test Factor

A comparison of the Z_i^* factors and P (Z_i^*) for the three studies (decreasing values of judgment factor 1.0 to 0.1 to 0.01) indicates that some reduction in probability

confidence level accompanies the decrease in factor numerical value. The physical interpretation of this judgement factor is bound into a designer - evaluator's attempt to establish a realistic not just feasible level of dependency of an integrated parameter (such as y - our effluent system rate) upon the selected set of measureable variables x_i . Since the set $\{x_i\}$ in this first study are laboratory measurable variables found from samples of the fluid of our waste-water - processing plant, their influence individually upon effluent should be expected to be less than 10%. The data ensemble for 0.1 has one contribution for x_6 approaching our 10% limitation. We will employ our $\mathcal{Z}_i^*$ test in a later adaptivity model evaluation.

The selection policy employed in determining parameters (i=1, 2, 6, 12, 13, 14, 15, and 16) and their corresponding identifications (φ = 18, 19, 20, 21, 22, 23, 24 and 25). An original set $\{x_i\}$ for (i = 1, 2, ---, 26) were found to have single linear correlation coefficients less than 0.10 for a number of those in the set. These were dropped in the curvilinear regression models. The reverse Markov process was adopted for the search for $f_i^{(2)}(x_i)$ rather than setting $f_i^{(2)}(x_i)=(b_i)x_i + (b_i)'(x_i^2)$ (1.27) where the b_i and $(b_i)'$ found as the coefficients or dependency sensitivity values falling out of the regression iteration process. The $\mathcal{Z}_i^*$ test is closely tied in with this Markov chain subprocess.

This test compares the linear model contribution for $x_i = \overline{x}_i$ to y with the Markov resultant contribution for the x_i variable.

Some difficulty was experienced in formulating the detail software programming for the several regression analyses. This possibly arose because of the nature of the programming which assigns some internal decision-making to the computer through application of an F and student t tests. As we shall learn from study of the trade-off two-parameter dependency plots many of the parameters exhibit parabolic modeling and low numerical values for linear correlation coefficients.

An Adaptivity Model

For a physical system under observation and evaluation, we seek confidence in our regression model (1.07) and are concerned about its utility under widely-varying J condition. For a study these sets may be M in number.

When we consider multiple bounding conditions:

$$\bar{y}_J \rightarrow w_J \qquad (J=1, 2, ---, M) \qquad (1.25)$$

$\{w_J\}$ is our true system major model variable set of values and

$\{x_{kJ}\}^{\circ}$ the system (k=1, 2, --- , n) parameter operating value set. A linear regression model for the J condition set is

$$\bar{y}_J = w_J = (b_o)_J + \sum_{k=1}^{n} (b_{kJ}) \cdot (\bar{x}_{kJ}) + \text{non-linear and correction terms} \qquad (1.26)$$

$$= (b_o)^{\circ}_J + \sum_{k=1}^{n} (b_{kJ})^{\circ} \cdot (\bar{x}_{kJ})^{\circ} \qquad (1.27)$$

for the selected and restricted design-operating value set $\{x_{kJ}\}^{\circ}$

The zero superscript identifies the values found for the condition of balance of (1.26), our (1.27) where the non-linear and correction terms disappear. Reference is made to "Statistical Methods in Engineering Experiments" by E. M. Bartee, published by Merrill Books, Col. Ohio, 1966.

We begin with our experimentally acquired data and seek the specialized condition set. This condition set satisfies a regression model (found via the Method of Least Squares) and a system constraint--derivable from (1.27). The constraint relationship is basically a statement of excessive or deficient capacity potential in value for the integrated dependent variable (w_J) for the $\underline{J}$ set of operating and environmental conditions. To formulate this constraint we rewrite (1.27) as

$$W_J = (w_J - b_{OJ}) \leq \sum_{k=1}^{N} (b_k)_J \cdot (x_k)_J \qquad (1.28)$$

$$= \sum_{k=1}^{n} (b_{kJ})^{o} \cdot (x_{kJ})^{o} \qquad (1.29)$$

for $n < N$

where variables $k=n+1, n+2, ---, N$ are of negligible influence (ie $b_{kJ} < \epsilon$ small) or they are fixed and included in the threshold value (b_{OJ}).

Let us begin with a multiple parameter regression analysis constrained to the $\underline{J}$ set of conditions and influenced by our desired optimal condition given by (1.29). For the $\underline{J}$ condition set we define our error of estimation:

$$(e_j)_J = (y_j - y')_J \qquad (1.30)$$

$$(\ j = 1, 2, \text{----}, m)$$

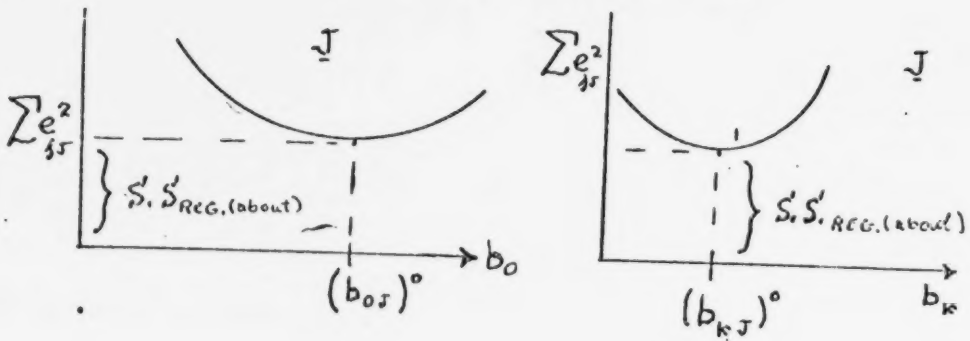
y' a point on the regression curve and

$$\sum_{j=1}^{m} (e_j)_J^2 = \sum_{j=1}^{m} \left[y_j - b_{kJ}^{T} \cdot x_{kJ} \right]^2 \qquad (1.31)$$

The coefficient set $(b_k)_J$ and its transpose $(b_k)_J^T$ are viewed as parameters rather than variables and are characteristic of our particular $\underline{J}$ regression model. We seek to minimize $\sum_{j=1}^{m} (e_j)_J^2$ with respect to each and every member of the set $(b_k)_J$. and to the same minimal value $\sum_{j=1}^{m} (e_j)_J^2$ min. value = sum of squares about regression

(1.31)

A parametric set of portraits would include:



for which the ordinate value S.S. about regression would have the same numerical value at each and all of the coordinates of the b; parameters found for the minimum of the least squares parameter.

Hence we set $$\frac{\partial}{\partial b_k}\left\{\sum_{j=1}^{m}(e_j)^2_s\right\} = 0 \qquad (1.32)$$

for k = 0,1, 2, ---, n

as parameter identification.

This yields the Jth system condition set of simultaneous equations:

$$\begin{aligned} b_1C_{11} + b_2C_{21} + \text{---} + b_nC_{n1} &= C_{y1} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ b_1C_{1n} + b_2C_{2n} + \cdots + b_nC_{nn} &= C_{yn} \end{aligned} \qquad (1.33)$$

where $$\left[C_{ik}\right]_J = \sum_{j=1}^{m} \left\{(x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k)\right\}_J \qquad (1.34)$$

$$\left[C_{yk}\right]_J = \sum_{j=1}^{m} \left\{(y_j - \bar{y})(x_{jk} - \bar{x}_k)\right\}_J \qquad (1.35)$$

or in matrix notation

$$\underline{b}_k^T \cdot \underline{C}_{ik} = \underline{C}_{yi} \qquad (1.36)$$

and $$(b_o)_J = (\bar{y})_J - \sum_{k=1}^{n} (b_k)_J^T \cdot (x_k)_J^o \cdot \qquad (1.37)$$

Equation (1.37) is our guarantee or anchor-point statement that the spatial point

$(\bar{y}, \bar{x}_k)_J$ for all k

is on the curvilinear regression model. Our minimal value for

$$\sum_{j=1}^{m} (e_j)_J^2 = S.S._{error}$$

$$= C_{yy} - \sum_{i,k=1}^{n} b_k C_{yk} \qquad (1.38)$$

for linear or parabolic variables. (p in number).

$$= C_{yy} - \sum_{i=1}^{p} b_i C_{yi}$$

$$W_J = \omega_J - (b_{oJ})^\circ \leq \sum_{k=1}^{n} (b_{kJ})^\circ \cdot (\bar{x}_{kJ})^\circ \qquad (1.39)$$

and we seek the design-operating values $\{x_{kJ}\}^\circ$

such that $$\Omega(x_k) = \sum_{J=1}^{M} W_J = \sum_{k=1}^{n} (x_k)^\circ \sum_{J=1}^{M} A_{Jk} = \sum_{k=1}^{n} A_k \cdot (x_k)^\circ \qquad (1.40)$$

The $\{x_k\}_J^\circ$ are the diagonal terms of the nxm matrix of $[x_{kJ}]$ for $k = J$. The complete set is found only for $n = m$. $x_{kk} \approx x_k^\circ$

J \ i					
J = 1	x_{11}°	x_{21}	x_{31}	- - -	x_{n1}
J = 2	x_{12}	x_{22}°	x_{32}	- - -	x_{3n}
.	.	.	x_{33}°	.	.
.	.	.	.	.	.
.	.	.	.	.	.
J = m	x_{1n}	-	-	- - -	$x_{nn}^\circ = x_{mm}^\circ$

To establish a system constraint to assist in our search for the set $\{x_{kJ}\}^\circ$ we select

$$x_{jk} = c_{jkJ} \cdot (x_{kJ})^\circ \qquad (1.41)$$

We assume that $(x_{kJ})^{\circ} = (\bar{x}_k)_J$ and $(x_i)_J = (x_{iJ})^{\circ}$

Relationships (1.32) and (1.33) become:

$$(C_{1K})_J = \bar{x}_1 \sum_{j=1}^{m} f_{j1} f_{jK} x^{\circ}_{KJ}$$

$$= (\bar{x}_1 x^{\circ}_{KJ} C_1 C_K \cdot m)_J \qquad (1.42)$$

for $\overline{C_1 C_K m}$ $\qquad = \sum_{j=1}^{m} f_{j1} f_{jK}$ $\qquad (1.43)$

= average numerical number $\overline{C_1 C_k}$ times the number of terms

= integrated value

and $(C_{yK})_J = (x_K)^{\circ}_J \cdot \sum_{j=1}^{m} f_{jK}(y_j - \bar{y})$ $\qquad (1.44)$

Our n(or m) simultaneous set of equations (1.33) becomes:

$$(b_1 \bar{x}_1 \cdot \bar{x}^{\circ}_{KJ} C_1)_J + (b_2 \bar{x}_2 \bar{x}^{\circ}_{KJ} C_2)_J + \cdots + (b_n \bar{x}_n \bar{x}^{\circ}_{KJ} C_n)_J$$

$$= (x^{\circ}_{KJ}) \cdot \sum_{j=1}^{m} f_{jK}(y_j - \bar{y}) \qquad (1.45)$$

for the Kth equation.

Rewriting it, we find:

$$(b_1 \bar{x}_1 \sum_{j=1}^{m} f_{j1})_J + \cdots + (b_n \bar{x}_n \sum_{j=1}^{m} f_{jn})_J$$

$$= \underline{\sum_{j=1}^{m} f_{jK}(y_j - \bar{y})} \qquad (1.46)$$

The right side of (1.46) is a weighted average value for

$$(\bar{y}-\bar{y}) \text{ and equals } (b_1\bar{x}_1C_1) \quad + \text{---+}(b_n\bar{x}_nC_n) \tag{1.47}$$

We can reform (1.47) into the linear model of y as a function of the arithmetic mean values for (x_k) .

$$(\dot{y})_J = (\bar{y})_J \; + \left\{ C_1^{\bullet}\bar{x}_1 + C_2^{\bullet}\bar{x}_2 + \text{---} + C_n^{\bullet}\bar{x}_n \right\}_J \tag{1.48}$$

Our equation set (1.43) has a redundancy of (m-1) so that (1.48) is our single constraint. For each J set of operating conditions (1.48) defines a model tangency curve to our curvilinear regression model. Only this tangency model satisfies our selection policy (1.41). At the tangency coordinates $(y,\bar{x}_i)_J$, (1.48) shares these with each of the trade-off plots (regression two-parameter dependency portraits)

$$(y)_J = (\bar{y}+B_0^{\bullet})_J \; + (C_i^{\bullet}x_i)_J \tag{1.49}$$

where $$B_0^{\bullet} = \sum_{\substack{R=1 \\ R\neq i}}^{n} C_R^{\bullet}(\bar{x}_R) \tag{1.50}$$

The slope characteristic for the constraint model is $(C_i^{\bullet})$ This must equal the slope of the curvilinear regression

model at the coordinate point $(y, \bar{x}_i)$. Referring back through our development of (1.49) we see that

$$(C^*)_J = (b_i C_i)_J = \left(b_i \cdot \sum_{j=1}^{m} r_{ji}\right)_J$$

$$= \left\{ b_i \cdot \sum_{j=1}^{m} \frac{r_{kj}}{r_{kk}} \right\}_J \qquad (1.51)$$

$$= \{b_i\}_J \qquad (1.52)$$

This must equal in magnitude and polarity the slope of our curvilinear regression model

$$(y)_J = (B_o)_J + \left\{ \sum_{\substack{R=1 \\ R \neq i}}^{n} f_R^{(2)}(x_R) + f_i^{(2)}(x_i) \right\}_J \qquad (1.53)$$

$$= \text{constant} + \left\{ f_i^{(2)}(x_i) \right\}_J \qquad (1.54)$$

$$\text{Slope} = \left(\frac{\partial y}{\partial x_i}\right)_{J, (\bar{x}_i)_J^\circ} = \left\{ \frac{\partial}{\partial x_i} \left(f_i^{(2)}(x_i) \right)_J \right\}_{(\bar{x}_i)_J^\circ} \qquad (1.55)$$

for $f_i^{(2)}(x_i) = b_i x_i + b_i^* x_i^2$ approximately (1.56)

$$\left(\frac{\partial f_i^{(2)}(x_i)}{\partial x_i}\right)_{J,(\bar{x}_i)_J} = \left\{b_i + 2 b_i^* \bar{x}_i\right\}_J$$

$$= \text{Constant} \qquad (1.57)$$

$$= (b_i')_J \longrightarrow (b_i)_J \qquad (1.58)$$

or b_i for a curve for which the non-linear term $b_i^* << b_i$ (1.59)

The linear regression model resulting from the processing of experimentally-derived sample parameter data provides a coefficient set $\{b_k\}_J$ and mean value set $\{\bar{x}_k\}_J$. Together these may be directly for system constraint used as definition for each J set of operating conditions. Moreover only one such constraint relationship exists for each J set. One also can provide for checking the system definition at each j measurement condition since:

$$w_j = y_j - \beta_{oj} = \sum_{k=1}^{n} \beta_{jk} a_{jk} \bar{x}_k + \text{non-linear terms}$$

for a total population. (1.60)

For a more defined set J of operating and environmental conditions

$$(w_j)_J = (y_j - \beta_{oj})_J = \sum_{k=1}^{n} \beta_{jk} \bar{x}_k + \sum_{k=1}^{n} \beta_{jk}(x_{jk} - \bar{x}_k) + \Delta_J \qquad (1.61)$$

where β_{jk} = sensitivity of the y variable measured at y_j due to changes in the x_k parameter.

$(b_0)_J$ = threshold value of y

$$= (\bar{y})_J - \sum_{k=1}^{n} (b_k)_J \cdot (\bar{x}_k)_J \qquad (1.62)$$

All of the required sets of coefficients are available from a computer-derived regression analysis. The $\{a_{jk}\}$ or $\{a_{jkJ}\}$ coefficient set are defined for each j measurement condition under J operating conditions and are identified as the ratio of the measured value of the $(x_k)_J$ parameter to its regression model (or J population) arithmetic mean value.

System Model Definition

We define a system model in terms of a limited number n of the total population of N recognizable parameters. (N-n) of the total parameter set may not be directly or indirectly measurable. Our limited set is practical for the j and J conditions. The usual system models are defined by the constraint sets:

$$\{w_j\} \gtrless \sum_{k=1}^{n} A_{jk}\, x_k \lesseqgtr \underline{A}_{jk} \cdot \{x_k\} \qquad (1.63)$$

$$= \sum_{k=1}^{n} A_{jk}\, x_k + \sum_{i=n+1}^{N} A_{ji} \cdot x_i \quad + \text{non-linear terms} \qquad (1.64)$$

$$x_i \geq 0 \qquad (k=1,2,\cdots,n)$$
$$(i=1,2,\cdots,N)$$
$$(j=1,2,\cdots,m)$$

$$\sum_{j=1}^{m} w_j = \left\{\sum_{j=1}^{m} A_{j1}\right\} x_1^{\circ} + \left\{\sum_{j=1}^{m} A_{j2}\right\} x_2^{\circ} + \cdots + \left\{\sum_{j=1}^{m} A_{jn}\right\} x_n^{\circ} \qquad (1.65)$$

$$W_J = \sum_{k=1}^{n} (b_k)_J \cdot (x_k)_J^{\circ} = \sum_{k=1}^{n} (b_k)_J \cdot \bar{x}_{kJ} \qquad (1.66)$$

W_J includes $\sum_{j=1}^{m} \beta_{oj}$

For the more general J set of system operating conditions:

$$\{W_J\} = \sum_{K=1}^{n} A_{JK}\,\bar{x}_{KJ} + \sum_{i=n+1}^{N'} A_{Ji}\,x_i \qquad \text{+ non-linear terms}$$

$$x_i \geq 0 \;;\; x_K \geq 0 \;;\; N' \neq N \qquad (1.67)$$

$$\sum_{J=1}^{M} W_J = \Omega(x_i) \qquad (1.68)$$

$$\sum_{J=1}^{M} W_J - \Omega_o(\bar{x}_K)^\circ = \sum_{K=1}^{n} (B_K)_J^\circ \cdot (\bar{x}_{KJ})^\circ - \Omega_o \qquad (1.69)$$

$$= \sum_{K=1}^{n} B_K^T \cdot \bar{x}_K - \Omega_o$$

$$= 0 \qquad (1.70)$$

and $$\left\{\bar{x}_{kJ}\right\}^\circ \in \left\{\bar{x}_k\right\}_J^\circ \qquad (1.71)$$

The empirical data may be used to verify that the experimentally-derived sensitivity factors are consistent with those derivable from mathematical optimization methodology. This is determined from the following:

At each J condition:

$$w_J = A_{J1}\,x_1 + A_{J2}\,x_2 + \cdots + A_{Jk}\,x_k + \Delta_J \qquad (1.72)$$

At the design-operating points for the $\{x_k\}$ we know that:

$$w_J = A_{J1}\,(x_{1J}^{\circ}) + A_{J2}\,(x_{2J}^{\circ}) + \cdots + A_{Jk}\,(x_{kJ}^{\circ}) \qquad (1.73)$$

At the $(J+1)$ condition

$$w_{(J+1)} = A_{(J+1),1}\,(\bar{x}_1 + \Delta x_1)_J + (A_{(J+2),2}(\bar{x}_2 + \Delta x_2) + \cdots + A_{(J+n),n}(\bar{x}_n + \Delta x_n) \qquad (1.74)$$

At the midway point $(J+\frac{1}{2})$, we find

$$(w_{(J+1)} - w_{(J)}) = \delta w_{(J+\frac{1}{2})} = A_{(J+1),1}\,\bar{x}_1 + A_{(J+1),1}(\Delta x_1) + \cdots A_{(J+1),n}\,\bar{x}_n + A_{(J+1),n}(\Delta x_n) - A_{J,1}\,\bar{x}_1 - - - A_{J,n}\,\bar{x}_n \qquad (1.75)$$

When we hold the change in conditions to small increments:

$$(\delta w)_{J+\frac{1}{2}} = (A_{(J+\frac{1}{2})})_1 (\Delta x_1)_J + (A_{(J+\frac{1}{2})})_2 (\Delta x_2)_J + \cdots + (A_{J+\frac{1}{2}})_n (\Delta x_n)_J \qquad (1.76)$$

and $(A_{(J+\frac{1}{2})})_i = \left(\frac{\partial w_{(J+\frac{1}{2})}}{\partial x_i}\right)\Bigg|_{(\bar{x}_i)_J}$ = sensitivity of w_J to the parameter x_i and measured at $(\bar{x}_i)_J$

$$\approx (A_J)_i \qquad (1.77)$$

Moreover, $(\delta_w)_{(J+\frac{1}{2})}$ represents the accumulative or "integrated" measure of the variation w variable seen as the sum of the weighted incremental changes for the limited set of system parameters. Likewise, if the $\{\Delta x_i\}$ are a set of permissable tolerances for the $\{x_i\}$ parameter set. $(\delta_w)_{(J+\frac{1}{2})}$ is the system variational or resultant predictable system tolerance for the (J+½) condition. In matrix array form our system is defined as:

Influence - Coefficient Matrix

System Conditions J \ i	$x_i \longrightarrow$					w_J	
J=1	A_{11}	A_{21}	A_{31}	...	A_{n1}	w_1	$w_{(J+\frac{1}{2})}$
J=2	A_{12}	A_{22}	A_{32}	...	A_{n2}	w_2	$=(w_2-w_1)$
.	.	.	.		.	.	
.	.	.	.		.	.	
.	.	.	.		.	.	
J=M	A_{1M}	A_{2M}	A_{3M}		A_{nM}	w_M	
	B_1	B_2	B_3	...	B_m	(x_i)	

$$\{B_i\}^T \cdot \{x_i\}^\circ = \bigcap (x_i^\circ) = \bigcap_0 (x_i^\circ) \qquad (1.78)$$

The design-operating matrix is

J \ i	$i \rightarrow$					$\Omega(x_{kJ}^{\circ})$
J=1	x_{11}°	x_{21}°	x_{31}°	...	x_{n1}°	$\Omega(x_{k1}^{\circ})$
J=2	x_{12}°	x_{22}°	x_{32}°	...	x_{n2}°	$\Omega(x_{k2}^{\circ})$
.	.	.	.		.	
.	.	.	.		.	
.	.	.	.		.	
J=M	x_{1M}°	x_{2M}°	x_{3M}°	...	x_{nM}°	$\Omega(x_{kM}^{\circ})$

Our selected tolerance vector $\{\Delta x_i\} = \begin{Bmatrix} \Delta x_1 \\ \Delta x_2 \\ . \\ . \\ . \\ \Delta x_n \end{Bmatrix}$

One form for the objective function seen as a pay-off variable is the following:

$$\Omega(x_i) = \underset{i=J}{\Omega(\bar{x}_{iJ})^{\circ}} - \sum_{J=1}^{M} A_{iJ} \underset{(i=J)}{\{\Phi_J(x_i) - \omega_J\}} \bar{x}_{iJ}^{\circ} \qquad (1.79)$$

The set $\left\{\bar{x}_{iJ}\right\}^{\circ}_{(i=J)}$ are the diagonal set of design-operating values for the array $\left\{x_{iJ}\right\}$. The $\Phi_J(x_i)$ or $\left\{f_i^{(2)}(x_i)\right\}_J$ are the system condition functions we described in the background discussion, page 7. They are defined by Taylor's series expansion through use of the trade-off plots and the compatible coordinate data. The major use for $\Omega(x_i)$ or its components $\Omega(x_{kJ}^{\circ})$ serve as a selection function for choice of the $\left\{x_{kJ}^{\circ}\right\}$ and for prediction for other than those J conditions measured or specified. Analysis of the Influence-Coefficient Matrix permits this solution.

Graphical Interpretation

Before we turn to a discussion of analytical procedures, let us form some graphical portraits to assist us in practical interpretation of our theoretical characterizations. From the sample sets of experimental data for our physical systems at a limited set of partially controlled conditions, we have performed a multiple parameter regression study. For the ith, and Kth parameter under the J condition set our portrait is

$$(B_o)'_J = (B_o)_J + (B_i \bar{x}_i + B_i'(\overline{x_i^2}))$$

$$(b_o)'_J = (b_o)_J + b_i \bar{x}_i - (b_i)'_2 (\overline{x_i^2}) - (b_i)'_1 \bar{x}_i$$

$$(B_o)_J = (B_o')_J + (B_K' - B_R)(\bar{x}_R) + (B_R)' \overline{x_R^2} = (b_o)_J + (b_i \bar{x}_i)_J - (b_R \overline{x_R})_J$$

$$(B_o)'_J = (b_o)'_J + (b_i \bar{x}_i)_J + \{(b_i)' \overline{x_i^2}\}_J - (b_R \bar{x}_R)_J - (b_R)'(\overline{x_R^2})_J$$

Figure I-1

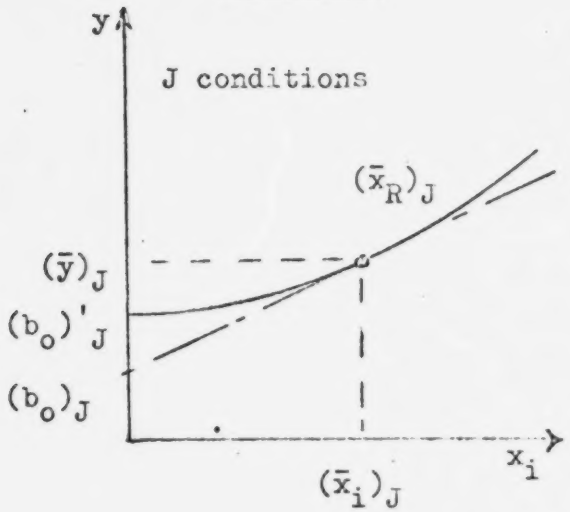
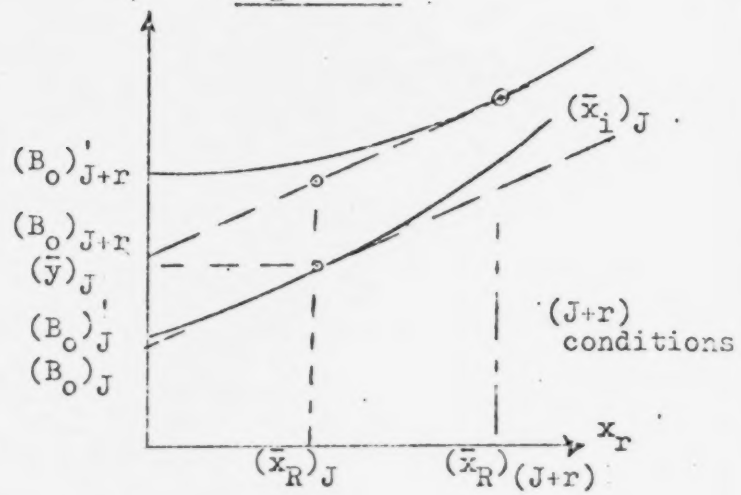


Figure I-2



$(b_R)_J = (b_R)_{J+r}$ = slopes at arithmetic means

In Figure-2, a plotting of the linear model for the (J+r) set of conditions shows a line of identical slope to the Jth regression line, with the same x_R but with different $(\bar{y})_{(J+r)}$ and intercept or threshold value. We interpret this portrait as one representing a system where the mean value x_R for parameter x_R is unchanged for conditions J and J+r. The sensitivity of (y) at J and J+r conditions which is equal to the slope of the linear model measured at x_R is also unchanged. However, a change in the threshold value from $(b_o)_R$ to $(b_o)_j$ and $(\bar{y})_J$ to $(\bar{y})_{J+r}$ indicates that some x_R do shift with change in operating conditions and the decrease shows these shifts are to lower values for positive parameter sensitivities and to higher values for those with negative sensitivities.

The linear regression model, when shown as two dimensional portraits of y variable against any x_i for several sets of J conditions are identical with our system trade-off plots where $(\bar{x}_i)_J = (x_i)_J^{\circ}$. This last relationship shows our identification or equating of sample-derived mean value with system-selected design-operating value for the Jth set of conditions. A typical non-linear portrait for a system trade-off study is shown in Figure I-3

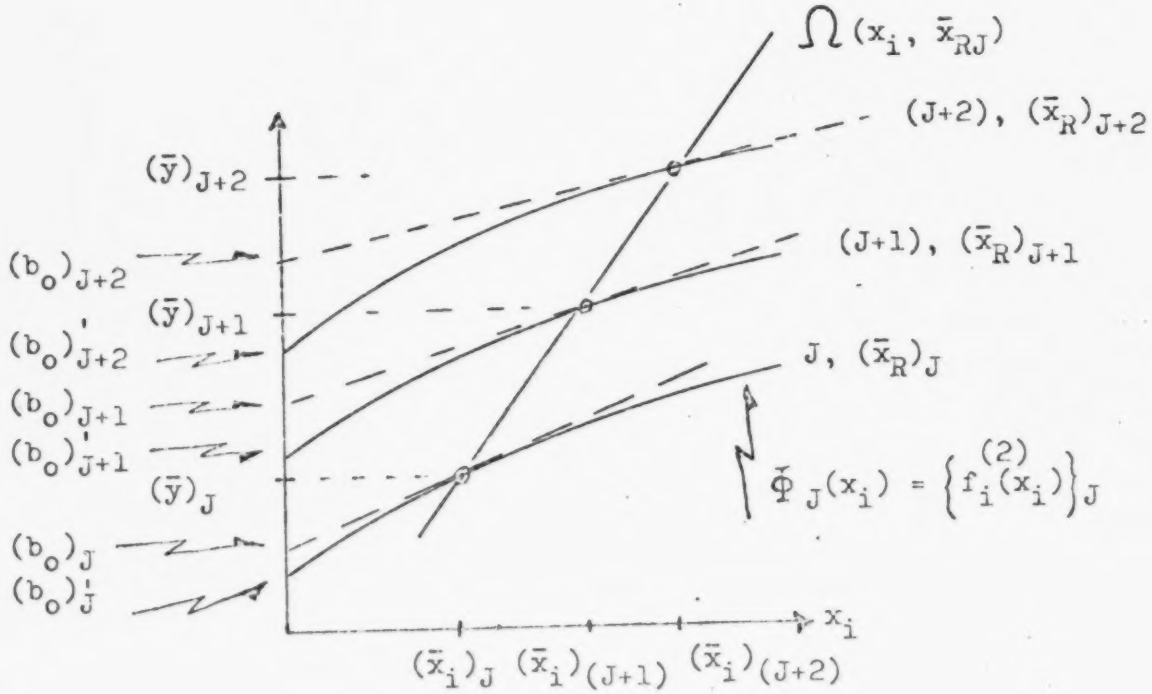


Figure I-3

$$(\text{slope})_J = \left(\frac{\partial y}{\partial x_i}\right)_{J,\ (\bar{x}_R)_J}$$, may be equal to linear regression model slope at $(\bar{y},\ \bar{x}_i,\ \bar{x}_R)_J$

$$(\text{slope})_{J+1} = \left(\frac{\partial y}{\partial x_i}\right)_{J+1,\ (x_R)_{J+1}}$$

$$(\text{slope})_{J+2} = \left(\frac{\partial y}{\partial x_i}\right)_{J+2,\ (x_R)_{J+2}}$$

If we restrict our system operation to small tolerance bands for all $\left\{x_i\right\}_J$ around $\left\{\bar{x}_i\right\}$, the trade-off $(\text{slopes})_J$ may be used as identical to the linear regression model slopes. Most systems analysis techniques restrict the system condition functions as shown in Figure I-3 to the characteristics of convex analytical functions. The further identification of the design-operating or central statistical values for the operating points as the singularity points in a Taylor's Series expansions for the systems conditions functions of $(y)_J = \left[f_i^2\ (x_i)\right]_J$ permits the utilization of variational techniques. For most system evaluation or design studies the slopes of the

$$\left\{ f_i(x_i) \,\middle|\, \bar{x}_R, \bar{x}_i, i \neq R \right\}_J$$ are more important

than the threshold values of (y_J) for $x_i \to 0$, $\bar{x}_R$ for $R \neq i$. Hence, the statistical regression models based upon linear dependency gives the required slopes at the $\left\{\bar{x}_i,\ \bar{y}_J\right\}_J$

coordinates -- the most wanted data.

Where one has ready access to a digital computer of sufficient capacity, a curvilinear multiple parameter regression analysis is practical.

Problem Solution Procedure via Statistical Optimization

A recommended orderly plan for the study and eventual solution of a system evaluation or design definition is desirable. If an early survey of the system indicates more parameters than ten in number per sub-system, a relative ordering or ranking of the most significant is needed to reduce the number to one convenient for analysis. A quick linear regression analysis for samples of ten measurements per parameter will give the slope values as measures of relative dependency of the system variable upon each parameter. Those of small dependency can be ignored with little influence upon the correctness or adequacy of our system model. Rarely, does one find more than eight to ten parameters of major significance or influence on system performance and operating characteristics.

A more careful multiple regression analysis will provide the required data including:

1. Sensitivities at each j condition and $\overline{x_i}$
2. Threshold values.
3. Correlation coefficients
4. Covariance

5. Mean or central values $\{\bar{x}_k\}$, $\{\bar{y}\}_J$

6. Deviation coefficients and $\{a_{ij}\}$, $\{\beta_{jk}\}_J$,

$\{b_k\}_J$, $\{A_{Jk}\}$, $\{A_{Jk}\}$, $\{B_{kJ}\}$

Utilizing the data found from the regression analysis formulate the system condition functions $\{f_i^2(x_i)\}_J$ as trade-off plots since we know the coordinate sets:

$$\left\{ \hat{y}_J \, , \, (\bar{x}_i)_J \, , \quad \left(\frac{\partial W_J}{\partial x_i}\right)_J , (\bar{x}_i)_J \right\}$$

From the Taylor's series expansions for the system condition functions $\{f_i^{(2)}(x_i)\}_J$ from the trade-off plots and the compatible coordinate data. If one desires the capability for prediction of system performance at other than the specific measured J sets of conditions, an optimal pay-off function is required. Equation (1.79) provides this statement.

This last step may be beyond that necessary for a particular study of interest. It is worthwhile to know that anlaytical techniques do exist for accomplishing this. A companion study -- " Trade-off Portraits and Design Charts" -- reports specific results for our waste-water processing plant.

Acknowledgement

The author gratefully has received valuable assistance from graduate students-Harry Frietchen and John Hungerford and from the staff of the Ohio University Computer Center in carrying the very extensive software programming and resultant data processing to practicality.

REFERENCES:

A. Bartee, E. M. : " Statistical Methods in Engineering Experiments", Merrill Books, Cols.O. -1966.

B. Gamble, E. H. : " The Modern Engineering Method", Vol 1, part 2, Purdue University, School of Mechanical Engineering, - 1963.

C. Feller, William : " An Introduction to Probability Theory and its Applications", Vols. I and II. Wiley, 1966.

D. Gamble, E. H. : " A Critque on the Design of Complex Systems", copyrighted June 1967, Library of Congress Card No. A 927900

STUDY B

"TRADE-OFF PORTRAITS AND DESIGN CHARTS"

(APPLIED TO DEMONSTRATION AND ANALYSIS OF THE SOMERSET OXIDATION-DITCH METHOD FOR WASTE-WATER PROCESSING)

BY E. H. GAMBLE, PH.D.
PROFESSOR OF SYSTEMS ENGINEERING
OHIO UNIVERSITY
ATHENS, OHIO

TRADE-OFF PORTRAITS AND DESIGN CHARTS
(ANALYTICAL TECHNIQUES)

I. INTRODUCTION

A. BACKGROUND

FOR SOMEWHAT LESS THAN TWO YEARS A SYSTEM DEMONSTRATION PROJECT FOR THE OXIDATION -DITCH WASTE-WATER PROCESSING PLANT LOCATED AT SOMERSET,OHIO HAS BEEN SPONSORED BY THE FEDERAL WATER POLLUTION CONTROL ADMINSTRATION OF HEW. THIS LARGE COMPLEX SYSTEM IS DEFINABLE IN TERMS OF FOUR INTERDEPENDENT SUB-SYSTEMS. THE MOST SIGNIFICANT BLOCK TO THEORETICAL SYSTEMS ANALYSES FOR BOTH IDENTIFICATION - DEFINITION AND EVALUATION IS SEEN TO BE THE DIFFICULTTY IN MAKING DIRECT VARIABLE-PARAMETERS MEASUREMENTS. FORTUNATELY, EMPIRICAL EXPERIMENTLY-DERIVED SAMPLE DATA FOR OUR STUDY SYSTEMMAY BE ACQUIRED UNDER VARIABLE AND NEARLY-CONTROLLABLE OPERATING AND IDENTIFIABLE AMBIENT ENVIRONMENTAL CONDITIONS. WITH A LARGE WEALTH OF LABORATORY INDIRECT ANALYSIS AND DIRECT PLANT MEASUREMENT DATA, THE DESIGNER-EVALUATOR IS FACED WITH A LARGE-SCALE DATA-PROCESSING AND DETECTIVE JOB.

OUR PRIMARY GOAL IS THE ENGINEERING DEFINITION OF PARAMETER DESIGN CHARTS. A DESIGN CHART MAY BE DESCRIBED AS A PORTRAIT OF AN IMPORTANT OUTPUT (SYSTEM CHARACTERISTIC IN HARDWARE NOMENCLATURE AND DIMENSIONS) VARIABLE AND ITS DEPENDENCY UPON A HARDWARE CHOICE VARIABLE. THIS HARDWARE SELECTED PARAMETER MUST BE DEFINED IN TERMS OF QUANTIFIED LIMIT VALUES. BECAUSE OUR PARTICULAR SYSTEM PARAMETER IS ALSO DEPENDENT UPON OTHER CHOICE VARIABLES, THESE MUST BE HELD AT DEFINED SETS OF FIXED VALUES. THE RESULTANT DESIGN PORTRAIT GIVES A FAMILY OF PLOTS. IF SUCH A CHART IS TO BE OF PRACTICAL DESIGN UTILITY, PREDICTION MUST BE MADE POSSIBLE AND EASY. ONLY AN ADDITIONAL CURVE INTERCONNECTING THE SELECTED

DESIGN-OPERATING CHOICE VALUES PROVIDES FOR VALUES OF THE SYSTEM VARIABLE AT OPERATING CONDITIONS INTERMEDIATE TO THOSE DEFINED BY THE FAMILY PLOTS ON THE CHART.

THE QUANTIFIED DESIGN CHART IS A COUSIN TO THE TRADE-OFF STUDY PORTRAIT USED BY ENGINEERS AS A PRELIMINARY STEP IN THEIR DECISION-MAKING PROCESS. THE TRADE-OFF STUDY MAY OR MAY NOT BE QUANTIFIED IN HARDWARE DIMENSIONAL UNITS AND IS VIEWED AS A THEORETICAL DEPENDENCY PORTRAIT. FOR A MULTIPLE PARAMETER PROBLEM MANY TRADE-OFFSTUDIES MUST BE MODIFIED UNTIL A CONSISTENT AND COMPATIBLE SET IS FOUND. USUALLY, AN ITERATIVE EMPIRICAL PROCESS OF GREAT LENGTH IS REQUIRED IN MODIFICATION, COMPARISON AND REMODIFICATION OF PLOTS --- HOPING TO CONVERGE UPON AN ACCEPTABLE CONSISTENT AND COMPATIBLE SET. TO AVOID THIS OFTEN-LUCKLESS PROCESS, AN ANALYTICAL METHOD HAS BEEN DEVELOPED. WE WILL REPORT ON A SIMPLIFIED EXAMPLE FOR ILLUSTRATION OF ITS APPLICATION. THE STATISTICAL TECHNIQUES DESCRIBED IN A COMPANION PAPER--- "STATISTICAL OPTIMIZATION"-HAVE SERVED TO FACILITATE THE THEORETICAL METHOD. IN ADDITION; THEY HAVE VERIFIED THE CLOSE RELATIONSHIP BETWEEN THE MULTIPLEREGRESSION MODELS OF APPLIED STATISTICS AND THE TRADE-OFF STUDIES. WITHOUT THE ASSIST OF THE MODIFIED CURVILINEAR REGRESSION TECHNIQUES, THE PURE ANALYTICAL METHOD WOULD HAVE SUFFERED. APPLICATION MAY HAVE BEEN LIMITED TO ONLY SMALL SYSTEMS AND SMALL NUMBER OF SETS OF DESIGN-OPERATING CONDITIONS.

B. Basic System Definition

The measurement set of data for each recognizable and identified variable (independent or near- independent) x_i and the companion dependent variable Y are related:

$$y_j = \beta_{0j} + \sum_{i=1}^{n} \beta_{ij}^{(1)} \cdot (x_i)_j + \sum_{i=1}^{n} (\beta_{ij})^{(2)} \cdot (x_i)^2_j$$

$$+ \text{ CORRECTION TERMS.} \qquad (2.01)$$

AND

$$w_j = (y_j - \beta_{0j}) = \sum_{i=1}^{n} \left[f_i^{(1)}(x_i) \right]_j + \sum_{i=1}^{n} \left[f_i^{(2)}(x_i) \right]$$

$$+ \text{ CORRECTION TERMS.} \qquad (2.02)$$

The $\left[f_i^{(1)}(x_i) \right]_j$ are viewed as the linear contributions of $(x_i)_j$ to y_j and the $\left[f_i^{(2)}(x_i) \right]_j$ as the parabolic contributions. We recognize (2.02) as an attempt to explain the quantified value y_j in terms of the selected contributors and the form of their contributions. Since a system exhibits most strongly an integrating functional characteristic , the population arithmetic means for the measurement subsets have more direct meaning . We rewrite (2.02) in the form :

$$W_J = \sum_{\substack{j=1 \\ J}}^{m} (y_j - \beta_{0j}) = \sum_{\substack{j=1 \\ J}}^{m} \left\{ \sum_{i=1}^{n} \left[f_i^{(1)}(x_i) \right]_j + \left[f_i^{(2)}(x_i) \right]_j \right.$$

$$+ \text{ CORRECTION FACTORS} \qquad (2.03)$$

$$W_J = (\bar{y}_j)_J - (b_0)_J = \sum_{i=1}^{n} \left[f_i^{(1)}(x_i) \right]_J + \left[f_i^{(2)}(x_i) \right]_J \qquad (2.04)$$

where the integrated correction factors are included in

THE THRESHOLD VALUE OF $(b_o)_J$.

THE $\underline{J}$ SUBSCRIPT IDENTIFIES THE PARTICULAR SET OF AMBIENT AND OPERATING CONDITIONS UNDER WHICH OUR SYSTEM IS VIEWED AND ANALYZED . THE CONTRIBUTIONS FUNCTIONS $\left[f_i^{(1)}(\rho_i)\right]_J$ AND $\left[f_i^{(2)}\right]_J$ POSSESS INTEGRATED VALUES BUT DO NOT REQUIRE THAT (x_i) AND (x_i^2) BE CONSTRAINED ONLY TO THEIR ARITHMETIC MEAN VALUES . MOREOVER, W_J IS CONSTRAINED TO ($\bar{Y}$)$_J$ THE $\underline{J}$ SUBSET POPULATION ARITHMETIC MEAN FOR THE DEPENDENT VARIABLE . LET US CONSIDER A SYSTEM WITH A PARTICULAR NUMBER OF DEPENDENT VARIABLES OF INTEREST : W_{KJ}

IF (2.04) DEFINES A MODEL FOR W_{KJ} FOR THE $\underline{J}$ SET OF CONDITIONS, THEN A TRADE-OFF PORTRAIT OF DEPENDENCY UPON A SINGLE VARIABLE $x_R = \rho_R$ IS DEFINED FOR SEVERAL $\underline{J}$ SETS, HENCE YIELDING FAMILY PLOTS . (2.04) IS REWRITTEN AS THE CONSTRAINT LINEAR CURVE : (tangency)

$$a_{RJK} \cdot (\rho_R)^\circ + \sum_{\substack{i=1 \\ i \neq R}}^{n} a_{iJK} \cdot (\bar{\rho}_i)_J = W_{JK} = \omega_{JK} - b_{oJK} \tag{2.05}$$

THE a_{fJK} ARE THE WEIGHTING OR INFLUENCE COEFFICIENTS AND FOR ρ_R :

$$(a_R)_{JK} = \frac{W_{JK} - \sum_{i=1}^{n} a_{iJK} \cdot (\bar{\rho}_i)_{JK}}{(\rho_R)_J^\circ} \tag{2.06}$$

FOR $i \neq R$, $(\bar{\rho}_i)_{JK}$ FIXED OR HELD AT THEIR SELECTED CENTRAL (DESIGN-OPERATING) OR $\underline{J}$ SET OF POPULATION ARITHMETIC MEAN VALUES.

THE BASIC TRADE-OFF PORTRAIT SHOWING THE DEPENDENCY OF THE (W_K) VARIABLE UNDER THE SEVERAL SELECTED J SETS OF CONDITIONS UPON (ρ_R) WITH THE OTHER VARIABLES HELD FIXED :

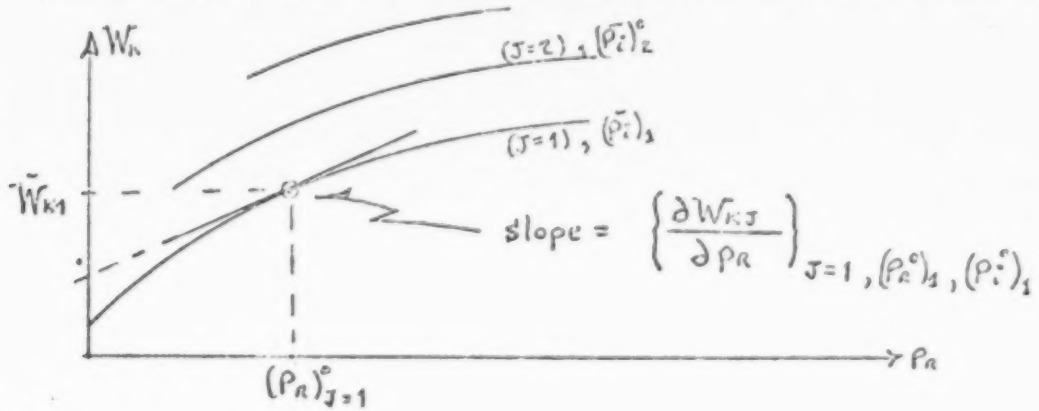


FIGURE II-1

THE VALUE FOR $(W_K)_{J=1} = \sum_{i=1}^{n} (a_{iK1})(\rho_i)^\circ_1 = W_{K1}$ (2.07)

IS THE COORDINATE OF THE TANGENCY LINE AT $\rho_R = (\rho_R)^\circ_{J=1}$ (2.08)

$$W_{K1} = (\bar{\mu})_1 - (b_o)_{J=1} \qquad (2.09)$$

THE CURVE OF W_{KJ} VERSUS ρ_R FOR THE (J=1) SET FOUND FROM (2.04) AS

$$(W_{KJ})_{\substack{J=1 \\ \rho_R \to 0}} = \mu_1 - b_{o1} = \sum_{\substack{i=1 \\ i \neq R}}^{n} \left[f_i^{(1)}(\bar{\rho}_i) + f_i^{(2)}(\bar{\rho}_i) \right]_{J=1} + \left[f_R^{(1)}(0) + f_R^{(2)}(0) \right]_{J=1} \qquad (2.10)$$

AND $(\rho_i)^\circ_J = (\bar{\pi}_i)_J$ (2.11)

WE IDENTIFY $\left[f_R^{(1)}(\rho_R) + f_R^{(2)}(\rho_R) \right]_J$ AS

$$\left[b_R^{(1)} \cdot \rho_R \right]_J + \left[b_R^{(2)} \cdot (\rho_R^2) \right]_J \qquad (2.12)$$

THE SLOPE AT THE POINT OF TANGENCY IS DEFINED AS:

$$\left\{ \frac{\partial W_{KJ}}{\partial \rho_R} \right\}_{J=1} = \left(b_R^{(1)} \right)_{J=1} + 2 \left(b_R^{(2)} \right)_{J=1} \cdot (\rho_R)^{\circ}_{J=1}$$

$$(\rho_R)^{\circ}_1 = \left(b_R^{(1)} \right)_1 + 2\, b_{R1}^{(2)} \cdot (\rho_R)^{\circ}_1 \qquad (2.13)$$

THE $(W_{KJ})_{(J=1)}$ VALUE FOR ($\rho_R \to 0$) FOR THE MODEL CURVE IS FOUND FROM (2.10) WHERE

$$\left[f_R^{(1)}(\rho_R) + f_R^{(2)}(\rho_R) \right]_{\rho_R \to 0} = 0 \qquad (2.14)$$

$$\left(W_{JK} \right)_{(J=1)} = y_1 - b_{01} = \sum_{i=1}^{n} \left[f_i^{(1)}(\bar{\rho}_i) + f_i^{(2)}(\bar{\rho}_i) \right]_{\substack{J=1 \\ i \neq R}} \qquad (2.15)$$

MODEL INTERCEPT VALUE

THE TANGENCY INTERCEPT LINE VALUE IS FOUND FROM (2.05)

FOR $\rho_R \to 0$

$$(W_{JK})_{(J=1)} = \sum_{\substack{i=1 \\ i \neq R}}^{n} b_{i1}^{(1)} \cdot (\rho_i)^{\circ}_{J=1} \qquad (2.16)$$

TANGENCY INTERCEPT VALUE

WHERE $W_{K1} > (W_{KJ})_{(J=1)}$ FOR $b_R^{(1)}$ IS POSITIVESLOPE,

TANGENCY INTERCEPT VALUE

AND SINCE W_{K1} INCLUDES AND EQUALS $(W_{KJ})_{(J=1)}$ (2.17)

TANGENCY INTERCEPT VALUE

$$+ b_{R1}^{(1)} \cdot (\rho_R)^{\circ}_1$$

FURTHER, IF $\left\{ W^{\circ}_{KJ},\ (\rho_R)^{\circ}_{J} \right\}$ AND $\left\{ W^{\circ}_{K,J+1},\ (\rho_R)^{\circ}_{(J+1)} \right\}$

ARE MEASUREABLE, ONE CAN FIND:

$$\left(\frac{W^{\circ}_{K,J+1} - W^{\circ}_{JK}}{(\rho_R^{\circ})_{J+1,K} - (\rho_R^{\circ})_{J,K}} \right) = \left(\frac{\partial W_{KJ}}{\partial \rho_R} \right)\Bigg|_{\substack{\rho^{\circ}_{\ell},\ W^{\circ}_{J,K} \\ @\ J+1/2}} \qquad (2.18)$$

IF THE (J) TO (J+1) CHANGE IS VIEWED AS SMALL COMPARED TO TOTAL CHANGE POSSIBLE, WE CAN DEFINE $\{a_{iJK}\}$ OR $\{b_{iJK}\}$ AND

$$-\left(\frac{\partial i}{\partial R}\right)_{JK} = \left(\frac{\partial \rho_R}{\partial \rho_i}\right)\Bigg|_{\substack{J,K \\ \bar{\rho}^{\circ}_{\ell},\ \bar{\rho}^{\circ}_{i}}} = \left(\frac{\partial W_J}{\partial \rho_i}\right)_K \Big/ \left(\frac{\partial W_J}{\partial \rho_R}\right)_K \quad \begin{matrix} \text{FIXED} \\ \text{FOR } J,\ \bar{\rho}^{\circ}_{R},\ \rho^{\circ}_{\ell},\ \bar{W}_J \end{matrix} \qquad (2.19)$$

WE EXPAND

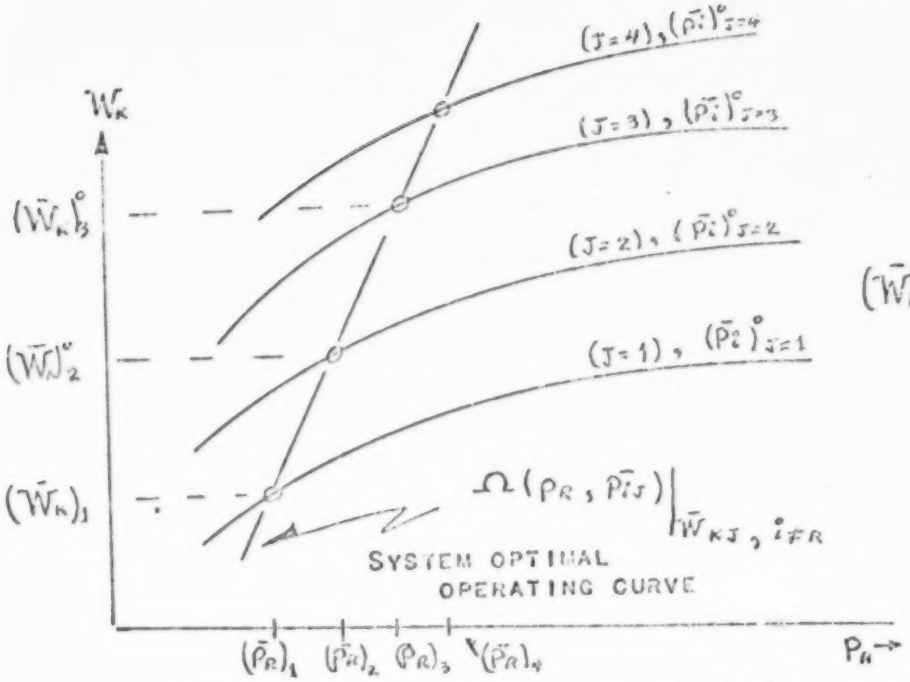
$$(W_K)_J = \sum_i (a_{iK})_J \cdot \rho_i + \sum_i \left(\frac{\partial W_K}{\partial \rho_i}\right)\Bigg|_{\substack{\bar{\rho}_i,\ J, \\ \bar{W}_{KJ}}} \cdot (\Delta \rho_i)$$

$$+ \sum_i \frac{\partial}{\partial \rho_i}\left(\frac{\partial W_K}{\partial \rho_i}\right)\Bigg|_{\substack{\bar{\rho}_i,\ J, \\ \bar{W}_{KJ}}} \cdot (\Delta \rho_i)^2$$

$$+ \sum_{\substack{i \\ i \neq R}} \frac{\partial}{\partial \rho_R}\left(\frac{\partial W_K}{\partial \rho_i}\right)\Bigg|_{\substack{\bar{\rho}_\ell,\ J \\ \bar{W}_{KJ}}} \cdot (\Delta \rho_i)(\Delta \rho_R) \qquad (2.20)$$

+ SMALL CORRECTION TERMS.

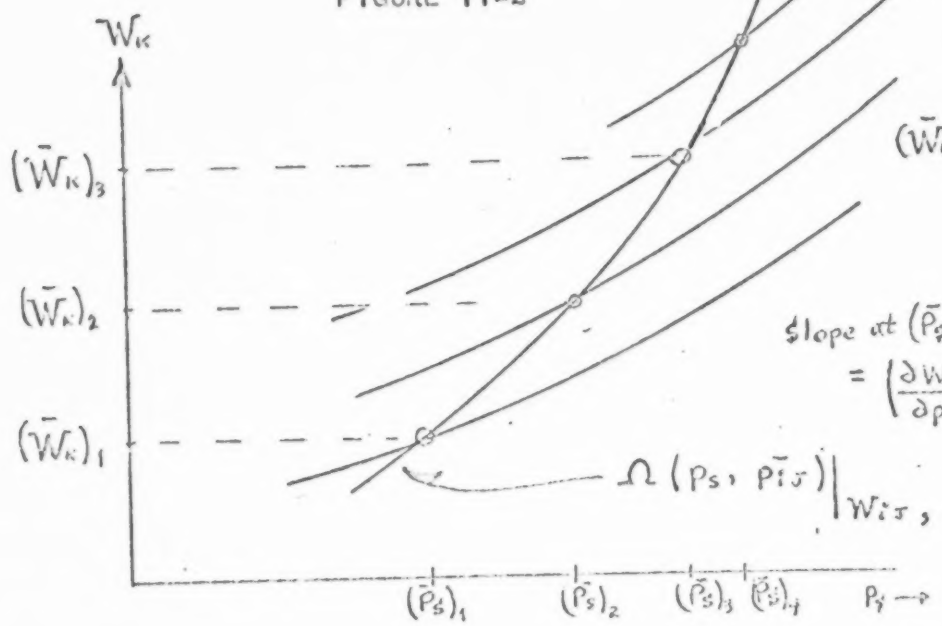
THERE WILL BE AN INDIVIDUAL TRADE-OFF PLOT FOR EACH OF OUR SELECTED SYSTEM INTERNAL AND OUTPUT PARAMETERS.



$$(\bar{W}_K)_1 = \sum_{i=1}^{n} a_{iK1} (\bar{P}_i)_1 \qquad (2.21)$$

" TRADE-OFF PORTRAIT FOR P_R VS. W_K "

FIGURE II-2



$$(\bar{W}_K)_2 = \sum_{i=1}^{n} a_{iK2} (\bar{P}_i)^\circ_2 \qquad (2.22)$$

$$\text{Slope at } (\bar{P}_S)_1, (\bar{W}_K)_1 = \left(\frac{\partial W_K}{\partial P_S}\right)_{K, J=1, \bar{P}_R, \bar{P}_S, \bar{P}_i} = \lambda_{S1}$$

FIGURE II-3
" SECOND TRADE-OFF PORTRAIT FOR W_K DEPENDENT UPON (P_S) "

WHERE THE a_{iJK} CELL VALUE MATRIX GIVES THE ARRAY OF INFLUENCE INTERDEPENDENCY COEFFICIENTS. OUR COMPUTING MATRIX IS :

J \ i	P_1	P_2	. . .	P_n	W_{KJ}
J = 1					W_{K1}
J = 2		a_{iJK}			W_{K2}
J = 3					W_{K3}
J = 4					W_{K4}
K	β_1	β_2	. . .	β_n	$\Sigma = \Omega$

FIGURE 11-4

FOR $\{P_i^\circ\}$ THE SELECTED OR FOUND DESIGN VALUES.

$$\beta_i = \sum_{J=1}^{M} a_{iJK} \qquad (2.23)$$

$$\sum_i \beta_i \cdot (p_i)^\circ = \Omega(p_i^\circ) = \sum_J \bar{W}_{KJ} = \sum_J \sum_i a_{iJK} \cdot p_i^\circ \qquad (2.24)$$

THE SENSITIVITY MATRIX

J \ i	(Δp_1) (Δp_2) . . . (Δp_n)	
J = 1/2	$\left(\frac{\partial W_{KJ}}{\partial p_i}\right) \cdot (\Delta p_i)$ $\bar{p}_{i\,1/2}$, $W_{K,\,J/2}$	$W_{K1} - W_{K0} = \Delta W_{K,\,J=1/2}$
J = 3/2		$W_{K2} - W_{K1} = \Delta W_{K,\,J=3/2}$
J = 5/2		$W_{K3} - W_{K2} = \Delta W_{K,\,J=5/2}$
J = 7/2		$W_{K4} - W_{K3} = \Delta W_{K,\,J=7/2}$

FIGURE 11-5
CELL-VALUE = SENSITIVITY, CHANGE IN W_{KJ} DUE ONLY TO P_i CHANGE

$\Omega(P_i)$ IS THE $\underline{N}$ - DIMENSIONAL SURFACE OF AN OBJECTIVE FUNCTION FOR SELECTION OF DESIGN - OPERATING VALUES $\{P_{iJ}\}^{\circ}$ FOR ALL $\underline{J}$ VALUES . IT HAS VALUE

$$\Omega(P_{i}^{\circ}) = \sum_{J} (\bar{W}_J)^{\circ} \qquad (2.25)$$

C. INTERPRETATION VIA STATISTICAL OPTIMIZATION

LET US INTERPRET THE (P_{ij}°) AS THE ARITHMETIC MEAN VALUES FOR $\{P_i\}$ UNDER THE GENERAL POPULATION CONDITIONS. THEN $\Omega(P_{i}^{\circ})$ IS THE VALUE OF THE SELECTION OR PAY-OFF FUNCTION AND OPTIMIZED FOR THE GENERAL POPULATION CONDITION.

FOR REDUCED SET, OR SUBSET

$\{P_{iJ}\}_{i \neq d}$ ARE THE $\underline{J}$ SUBSET OF DESIGN-OPERATING VALUES AND IDENTIFIABLE AS

$$\{\bar{P}_{\bar{i}J}\} = \{P_{iJ}\}^{\circ} \qquad (2.26)$$

OUR TWO- DIMENSIONAL PLOTS ARE THE $\underline{J}$ FAMILY PLOT REGRESSION MODELS FOR

$$(y)_J \text{ VS. } f_R^{(2)}(x_R),\ x_R \neq x_i \qquad (2.27)$$

OR

$$W_{JK} \text{ VS. } \Phi_J(P_R, P_{i\bar{J}}^{\circ}) \qquad (2.28)$$

FOR THE SEVERAL J'S.

THE TRADE-OFF SELECTION FUNCTION HAS FORM:

$$\Omega(P_R, P_{i\bar{J}}^{\circ}) \qquad (2.29)$$

WITH COMMON INTERSECTION COORDINATES:

$$(W_{JK})^{\circ} = (\bar{W}_{KJ})$$

$$P_{RJ}^{\circ} = P_{\bar{R}J} \qquad (2.30)$$

AND $(P_{iJ})^{\circ} = (P_{\bar{i}\bar{J}})$

AN EQUIVALENT SELECTION POLICY FOR THIS PARTICAL PLOT IS

$$\left.\frac{\partial \Omega(P_R, P_{i\bar{J}}^{\circ})}{\partial P_i}\right|_{\substack{\bar{P}_{RJ},\ J \\ \bar{P}_{iJ}}} = 0 \qquad (2.31)$$

11. An Illustrative Example of the Analytical Method

When our system modelfor multiple parameters and multiple conditions exceeds an (NXM) = (3x3) matrix array, one finds difficulty in forming proper interpretations quickly. To show the basis of the trade-off study method for design operating model definition, let us form a 3x3 problem and carry out the equivalent of a hand calculation. The following matrix has been givento us, or found via selection in some engineering way. Each cell value is (A_{IJK}), the influence coefficient:

J \ I	P	P	P	W_{KJ}
J=1	0.5	0.25	1.0	175
J=2	1.0	0.60	1.0	260
J=3	2.0	1.10	1.0	410
	3.5	1.95	3.0	845

FIGURE 11-6

We select P_3 as the dependency parameter and recognize $\{W_{KJ}\}$ as firm capacity potential values for our system under the J=1,2, and 3 set of operating and ambient conditions. The set of J constraint inequalities are :

$$\begin{aligned} 0.5\ P_1 + 0.25\ P_2 + P_3 &\leq 175 \\ 1.0\ P_1 + 0.60\ P_2 + P_3 &\leq 260 \\ 2.0\ P_1 + 1.1\ P_2 + P_3 &\leq 410 \end{aligned} \qquad (2.32)$$

and overall pay-off function for the population system:

$$3.5\ (P_{11}^{\circ}) + 1.95\ (P_{22}^{\circ}) + 3.0\ (P_{33}^{\circ}) = 845 = \Omega(P_{ii}^{\circ}) \qquad (2.33)$$

THE CORRESPONDING SYSTEM CONDITION FUNCTIONS $\Phi_J(P_1, P_2)$ EQUAL TO P_3 ARE RECOGNIZABLE AS THE $\left[f_i^{(2)}(x_i) \right]_J$ (2.34) CONTRIBUTIONS (FUNCTIONS) OF THE EQUIVALENT STATICAL REGRESSION MODEL DEFINITION (SEE 2.03 AND 2.04). OUR FUNCTIONAL SET IS :

$$\Phi_J(P_1, P_2) = P_3 = (P_{3J}^{\circ}) + \sum_{i=1}^{2} \lambda_{iJ} \cdot (P_i - P_{iJ}^{\circ}) - \frac{\Delta_J}{a_{3J}} \qquad (2.35)$$

THE λ_{iJ} ARE THE LAGRANGIAN PARAMETERS RELATING THE INFLUENCED DEPENDENCY OF Φ_J BY P_i. THE TERM $\frac{\Delta_J}{a_{3J}}$ CONTAINS THE NONLINEAR TERMS AND $\{W_J\}$ HAVE INCLUDED THE CONTRIBUTIONS OF ALL P_i FOR $i > 3$. SO THAT SYSTEM I MODEL IS THE 3x3 REDUCED (BUT CORRECTED) MODEL FOR A MUCH LARGER SYSTEM.

$$\Delta_J = \sum_{i=1}^{3} \frac{1}{2} \left(\frac{\partial^2 P_3}{\partial P_1^2} \right)\Bigg|_{(P_{iJ}^{\circ})} \cdot (P_1 - P_{iJ})^2 + \frac{1}{2!} \left(\frac{\partial^2 P_3}{\partial P_1 \partial P_2} \right)\Bigg|_{\substack{P_{1J}^{\circ} \\ \text{and} \\ P_{2J}^{\circ}}} \cdot (P_1 - P_{1J}^{\circ})(P_2 - P_{2J}^{\circ})$$

+ HIGHER ORDER TERMS.

THE λ_{iJ} MATRIX GIVES THE SLOPE VALUES FOR $\Phi_J(P_1, P_2)$ WHEN MEASURED AT THE $(P_{iJ})^{\circ}$, WHICH ARE TREATED AS TAYLOR SINGULARITY POINTS. THESE BECOME $(\bar{x}_{iJ}) = (x_{iJ})^{\circ}$, OUR SELECTED SET OF DESIGN- OPERATING POINTS FOR DESIGN CHART CONSTRUCTION.

$$\lambda_{iJ} = \begin{vmatrix} -\frac{a_{11}}{a_{31}} & -\frac{a_{21}}{a_{31}} & -\frac{a_{31}}{a_{31}} \\ -\frac{a_{12}}{a_{32}} & -\frac{a_{22}}{a_{32}} & -\frac{a_{32}}{a_{32}} \\ -\frac{a_{13}}{a_{33}} & -\frac{a_{23}}{a_{33}} & -\frac{a_{33}}{a_{33}} \end{vmatrix} = \begin{vmatrix} -0.5 & -0.25 & -1.0 \\ -1.0 & -0.6 & -1.0 \\ -2.0 & -1.10 & -1.0 \end{vmatrix} \qquad (2.37)$$

THE SET P_{1J} ARE THE OPTIMAL SOLUTIONS FOR THE SYSTEM WHICH UTILIZES THE FULL CAPACITY POTENTIALS.

$$W_{JK} = a_{1JK}\,(P_{1J})^{\circ} + a_{2JK}\,(P_{2J})^{\circ} + a_{3J}\,(P_{3J})^{\circ} \qquad (2.38)$$

TWO TRADE-OFF PLOTS ARE NEEDED FOR ILLUSTRATION. THE FAMILY PLOTS OF $\Phi_J(P_1, P_{2J}^{c})$ AND $\Phi_J(P_2, P_{1J}^{\circ})$. THE COMPLETED PLOTS ARE GIVEN ASFIGURES: PLOTSI AND II , RESPECTIVELY. IN ORDER TO FIND $\{P_{ii}\}^{\circ}$ AS THE REDUCED SET OF OPTIMAL AND GENERAL POPULATION MEAN VALUES $\{\bar{P}_i\}$ WE MUST DIAGONALIZE THE EQUALITY MATRIX :

$$|\lambda_{1J}| \cdot \begin{Bmatrix} P_{11} \\ P_{22} \\ P_{33} \end{Bmatrix} = \begin{Bmatrix} 175 \\ 260 \\ 410 \end{Bmatrix} \qquad (2.39)$$

THE SOLUTION OF WHICH IS $\{100\}$ OR $\bar{P}_1 = \bar{P}_2 = \bar{P}_3 = 100$ (2.40)

THE TWO TRADE-OFF PLOTS

CASE $P_3 = -\lambda_{11} \cdot P_1 - \lambda_{21} \cdot P_{21}^{\circ} + \dfrac{W_1 - \Delta_1}{a_{31}}$

$(J=1)$ $P_3 = -\lambda_{11} \cdot (P_{11}^{\circ}) - \lambda_{21} \cdot P_2 + \dfrac{W_1 - \Delta_1}{a_{31}}$ (2.41)

$(J=2)$ $P_3 = -\lambda_{12} \cdot P_1 - \lambda_{22} \cdot P_{22}^{\circ} + \dfrac{W_2 - \Delta_2}{a_{32}}$ (2.42)

$P_3 = -\lambda_{12} \cdot (P_{12}^{\circ}) - \lambda_{22} \cdot P_2 + \dfrac{W_2 - \Delta_2}{a_{32}}$ (2.43)

$(J=3)$ $P_3 = -\lambda_{13} \cdot P_1 - \lambda_{23} \cdot P_{23}^{\circ} + \dfrac{W_3 - \Delta_3}{a_{33}}$

$P_3 = -\lambda_{13} \cdot (P_{13}^{\circ}) - \lambda_{23} \cdot P_2 + \dfrac{W_3 - \Delta_3}{a_{33}}$

AN INSPECTION OF λ_{IJ} IN (2.37) SHOWS THAT THE SLOPE VALUES ARE ALL NEGATIVE AND THE Φ_J ARE CONCAVE DOWNWARD. WE HOPE TO RESTRICT THE Φ_J TO CONVEX FUNCTION CHARACTERISTICS AND RULE OUT EXTREMA FOR $0 < P_I < P^{\circ}_{IJ}$ TO AVOID MULTIPLE-VALUEDNESS.

A POSSIBLE FIRST SELECTION POLICY WOULD BE TO HAVE THE Φ_J MAXIMA OCCUR AT THE INTERCEPTS WITH THE P_3 ORDINATE AXIS. UNFORTUNATELY THIS POLICY LEADS TO A COMPLEX FORM (FOR THIS EXAMPLE) FOR THE CURVE INTERCONNECTING THE $(P_{IJ})^{\circ}$. THIS CURVE IS OUR $\Omega(P_1, P^{\circ}_{2J})$ OR $\Omega(P^{\circ}_{1J}, P_2)$, THE COMPANION PAY-OFF FUNCTIONS. WE MAKE USE OF (2.41), (2.42) AND (2.43) TOGETHER WITH SET (2.38) OR THE ALTERNATE FORMS OF (2.35). ANOTHER LOOK AT MATRIX (2.37) SUGGESTS THAT THE LEAST CHANGE OCCURS IN λ_{I2} FOR I= 1,2,3. FOR (2.33), CASE (J=2) WE FIND :

$$\Delta_2(P_{12}) = 260 - P^{\circ}_{12} - 0.6\,P^{\circ}_{22} - P^{\circ}_{32} = 0 \qquad (2.44)$$

SINCE $P^{\circ}_{22} = 100$ FROM (2.40).

$$\text{HENCE } P^{\circ}_{32} = 200 - P^{\circ}_{12} \qquad (2.45)$$

J = 2

FOR PLOT I , (J=2),

$$(P_3)_{\substack{\text{LIMIT}\\ \text{INTERCEPT}}} = 200 - 1.0P_1 - (\Delta_{12})_{@\ P_1 \to 0} \qquad (2.46)$$

$$= 200 - (\Delta_{12})_{P_1 \to 0} \qquad (2.47)$$

FOR PLOT II, (J=2)

$$(P_3)_{\substack{\text{LIMIT}\\ \text{INTERCEPT}}} = 260 - 1.0P^{\circ}_{12} - 0.60P_2 - (\Delta_{22})_{P_2 \to 0} \qquad (2.48)$$

$$= 260 - (\Delta_{22})_{P_2 \to 0} - P^{\circ}_{12} \qquad (2.49)$$

FOR THE SAME P_3 LIMIT VALUE :

$$P_3 = 200 - (\Delta_{12})_{P_1 \to 0} = 260 - (\Delta_{22})\Big|_{P_2 \to 0} - P_{12}^{\circ} \qquad (2.50)$$

OR $$P_{12}^{\circ} = 60 + (\Delta_{12})_{P_1 \to 0} - (\Delta_{22})\Big|_{P_2 \to 0} \qquad (2.51)$$

WE ARBITRARILY TAKE $P_{12}^{\circ} < P_{22}^{\circ} = 100$ AT VALUE 90.

THEN (2.51) YIELDS :

$$90 = 60 + (\Delta_{12})\Big|_{P_1 \to 0} - (\Delta_{22})\Big|_{P_2 \to 0} \qquad (2.52)$$

WHICH CAN BE SATISFIED BY

$$(\Delta_{12})\Big|_{P_1 \to 0} = 30 \qquad (\Delta_{22})\Big|_{P_2 \to 0} = 0 \qquad (2.53)$$

WITH $(P_3)_{\substack{\text{LIMIT} \\ \text{INTERCEPT}}} = 200 - 30 = 170$

OR

$= 260 - 90 = 170$

UTILIZING (2.35), WE FIND :

FOR PLOT II, $$(P_3)_{\text{LIMIT}} = P_{32}^{\circ} + 60 - \left(\frac{\Delta_{22}}{a_{32}}\right)\Big|_{P_2 \to 0} \qquad (2.54)$$

PLOT I, $$(P_3)_{\text{LIMIT}} = P_{32}^{\circ} + P_{12}^{\circ} - \left(\frac{\Delta_{12}}{a_{32}}\right)\Big|_{P_1 \to 0} \qquad (2.55)$$

$$= 200 - (\Delta_{12})\Big|_{P_1 \to 0} = 170 \qquad (2.56)$$

THEN : $P_{12}^{\circ} = 90$, $P_{32}^{\circ} = 110$ SATISFY (2.54) AND (2.55)

J=1 FROM (2.41) WE WRITE :

PLOT I $(P_3)_{LIMIT} = 175 - 0.5\, P_{11}^{\circ} - 0.25\, P_2 - (\Delta_{21})\big|_{P_2 \to 0}$ (2.57)

$= 125 - (\Delta_{21})\big|_{P_2 \to 0}$ (2.58)

PLOT II $(P_3)_{LIMIT} = 175 - 0.25\, P_{21}^{\circ} - 0.5 P_1 - (\Delta_{11})\big|_{P_1 \to 0}$ (2.59)

$= 175 - 0.25\, P_{21}^{\circ} - (\Delta_{11})\big|_{P_1 \to 0}$ (2.60)

OUR SYSTEM DESIGN MODEL (2.37) REQUIRES THAT :

$175 = 0.5\, P_{11}^{\circ} + 0.25\, P_{21}^{\circ} + P_{31}^{\circ}$ (2.61)

OR

$125 = 0.25\, P_{21}^{\circ} + P_{31}^{\circ}$ (2.62)

ALTERNATE (2.35) FORMULATION SAYS :

PLOT I $(P_3)_{LIMIT} = P_{31}^{\circ} + 0.25\, P_{21}^{\circ} - (\Delta_{21})\big|_{P_2 \to 0}$ (2.63)

PLOT II $(P_3)_{LIMIT} = P_{31}^{\circ} + 0.5\, P_{11}^{\circ} - (\Delta_{11})\big|_{P_1 \to 0}$ (2.64)

$= P_{31}^{\circ} + 50 - (\Delta_{11})\big|_{P_1 \to 0}$ (2.65)

SINCE $P_{33}^{\circ} = 100$, WE PREFER $P_{31}^{\circ} > P_{32}^{\circ} > P_{33}^{\circ}$ (2.66)

ARBITRARILY SELECT $P_{31}^{\circ} = 115$ WHICH IS PERMITTED BY (2.66), PROVIDED $100 < P_{32}^{\circ} < 115$. THEN (2.62) YIELDS:

$0.25\, P_{21}^{\circ} + 115 = 125$

$P_{21}^{\circ} \to 40$

AND (2.63) BECOMES:

$(P_3)_{LIMIT} = 115 + 10 - (\Delta_{21})\big|_{P_2 \to 0}$ (2.67)

$= 125 - (\Delta_{21})\big|_{P_2 \to 0}$

$(P_3)_{LIMIT} = 165 - (\Delta_{11})\big|_{P_1 \to 0}$ (2.68)

(2.67) AND (2.68) ARE SATISFIED BY

$(P_3)_{LIMIT} = 125,\ (\Delta_{21})\big|_{P_2 \to 0} = 0,\ (\Delta_{11})\big|_{P_1 \to 0} = 40$

J=3 OUR FIRST ESTIMATE FOR $(P_3)_{\substack{LIMIT \\ J=3}}$ WAS 255. LET US ADOPT THIS.

THEN THE TAYLOR'S EXPANSION (2.35) GIVES:

$$255 = P^{\circ}_{33} - 2.0\,(P_1 - P^{\circ}_{13}) - 1.1(P_2 - P^{\circ}_{23}) - \Delta_3 \qquad (2.69)$$

THE DESIGN RELATIONSHIPS (2.37) GIVE :

$$410 = 2.0\,P^{\circ}_{13} + 1.1P^{\circ}_{23} + 1.0P^{\circ}_{33}$$
$$310 = 2.0P^{\circ}_{13} + 1.1P^{\circ}_{23} \qquad (2.70)$$

PLOT I $\quad (P_3)_{LIMIT} = P^{\circ}_{33} + 1.1P^{\circ}_{23} - (\Delta_{23})\big|_{P_2 \to 0}$

$$= 255 - (\Delta_{23})\big|_{P_2 \to 0}$$

$$= 100 + 1.1P^{\circ}_{23} - (\Delta_{23})\big|_{P_2 \to 0} \qquad (2.71)$$

PLOT II $\quad (P_3)_{LIMIT} = 255 - (\Delta_{13})\big|_{P_1 \to 0} \qquad (2.72)$

$$= P^{\circ}_{33} + 2.0P^{\circ}_{13} - (\Delta_{13})\big|_{P_1 \to 0}$$

FOR $\quad P_{13} < P_{12} < P_{11}$

SELECT $\quad P_{13} = 90$.

WE WISH $\quad P^{\circ}_{21} < P^{\circ}_{22} < P^{\circ}_{23}$

OR $\quad P^{\circ}_{23} > 100$

OUR FIRST SET OF VALUES WAS $\begin{Bmatrix} 77.5 \\ 141 \\ 100 \end{Bmatrix}$

THIS YIELDS $\Delta_{23} = \Delta_{13} \to 0$ OR THE Φ_3 FUNCTIONS HAVE ZERO SLOPE AND MAXIMA AT $P_1 = 0$, $P_2 = 0$.

IF WE TRY TO LOWER THE $(P_3)_{LIMIT}$ TO 250, THE SIGN OF $(\Delta_{13})\big|_{P_1 \to 0}$ BECOMES NEGATIVE AND VIOLATES OUR CONVEX FUNCTION RESTRICTION.

WITH THE BASIC RELATIONSHIP :

$$410 = 2\,P^{\circ}_{13} + 1.1\,P^{\circ}_{23} + 1.0\,P^{\circ}_{33}$$

$$255 = \left(\Delta_{23}\right)\Big|_{P_2 \to 0} = 310 - 1.1\,P^{\circ}_{23}$$

$\left(\Delta_{23}\right)\Big|_{P_2 \to 0}$ MUST VANISH.

OUR SECOND ITERATION SET FOR $\left\{P_{IJ}\right\}^{\circ}$ MATRIX ARRAY IS :

	I=1	I=2	I=3
J=1	100	40	115
J=2	90	100	110
J=3	77.5	141	100

(2.73)

PLOT II BECOMES:

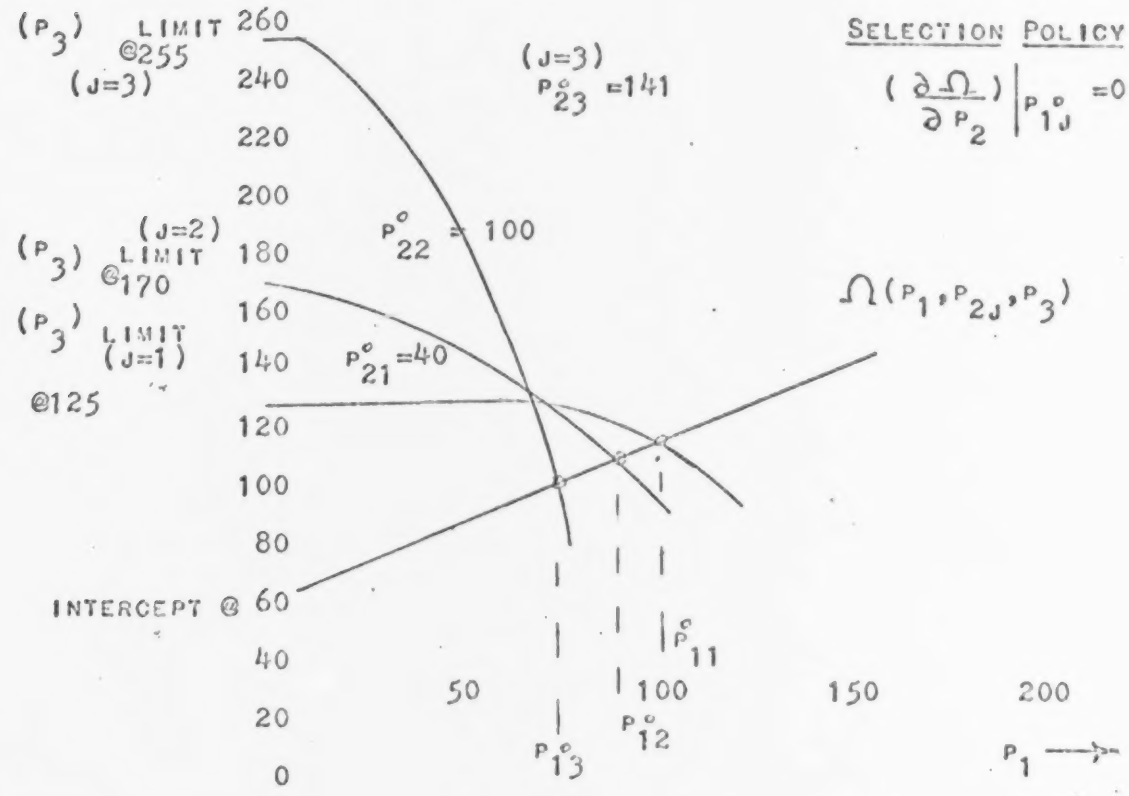


FIGURE 1.-7

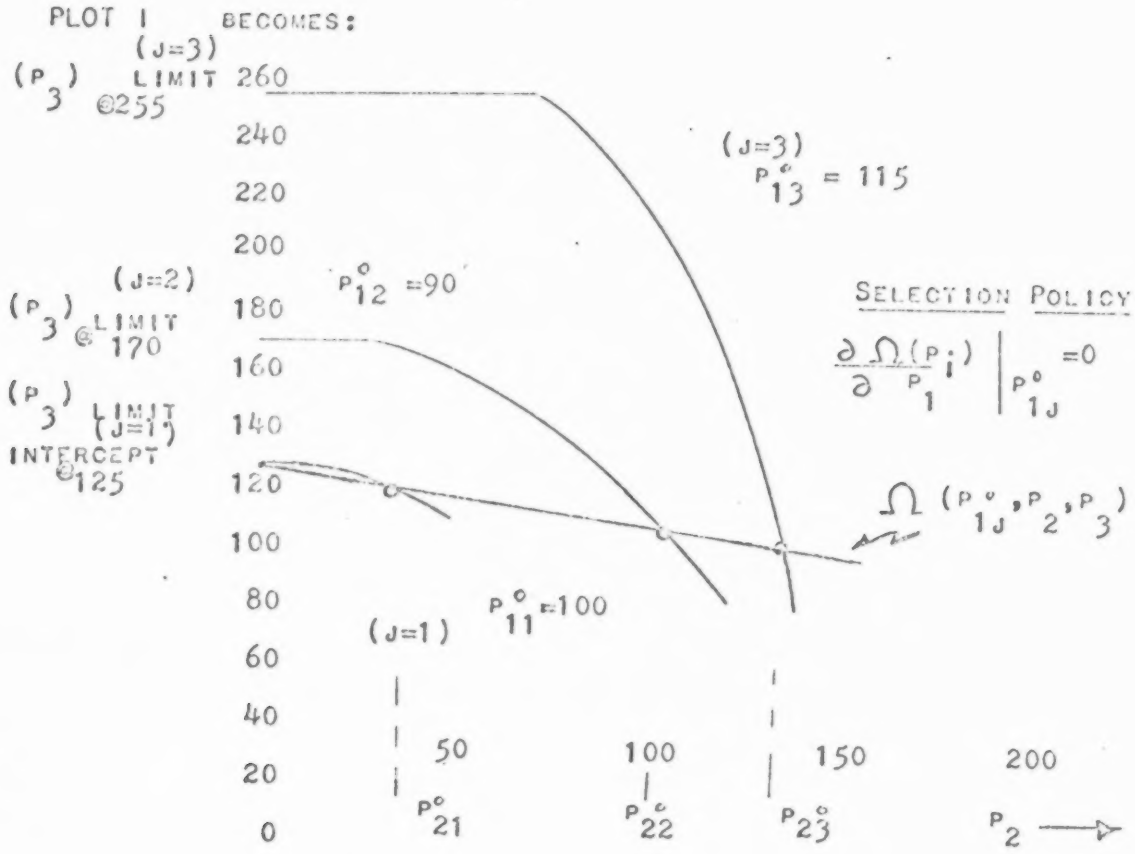


FIGURE II-8

FOR $\Omega(P_1, P_{2J}, P_3)$, $\{P_3 + 0.5P_1 = 60\}$ (2.74) IS EQUATION OF STRAIGHT LINE DERIVED FROM $\frac{\partial \Omega(P_1)}{\partial P_2}\Big|_{P^\circ_{21}} = 0$.

FOR $\Omega(P_{1J}, P_2, P_3)$, $\{P_3 - 0.25 P_2 - 125 = 0,\}$ (2.75) IS EQUATION OF STRAIGHT LINE DERIVED FROM

$$\frac{\partial \Omega(P_i)}{\partial P_1}\Big|_{P^\circ_{11}} = 0$$

OUR OBJECTIVE CRITERION OR PAY-OFF FUNCTION $\Omega(P_1) = \Omega(x_1)$ IS BY THE THEORY OF INEQUALTIES OR CONVEX FUNCTIONS THE OPTIMAL FUNCTION FORMED FROM THE DIAGONALIZATION OF THE $\{P_i\}$ VECTOR AND THE CONSTRAINT FUNCTIONS..

$$\gamma_i \cdot \varphi_J(P_1) = \Omega(P_1) \qquad \text{AND}$$

$$\Delta_J(P_1) = W_J - \sum_{i=1}^{3} a_{iJ}(P_i)$$

$$= \sum_{i=1}^{3} a_{iJ}\,(P_1 - P^{\circ}_{1J})$$

$$= a_{3J}\left[\phi_J(P_1, P_2) - P^{\circ}_{3J}\right]$$

$$+ a_{1J}\,(P_1 - P^{\circ}_{1J}) + a_{2J}\,(P_2 - P^{\circ}_{2J}) \qquad (2.76)$$

ALSO

$$\Omega(P_1) = \begin{vmatrix} \gamma_1 & 0 & 0 \\ 0 & \gamma_2 & 0 \\ 0 & 0 & \gamma_3 \end{vmatrix} \cdot \begin{Bmatrix} \varphi_1(P_1) \\ \varphi_2(P_1) \\ \varphi_3(P_1) \end{Bmatrix}$$

$$= \gamma_1 \frac{\Delta_1(P_1)}{\gamma_1} + \gamma_2 \frac{\Delta_2(P_1)}{\gamma_2} + \gamma_3 \cdot \frac{\Delta_3(P_1)}{\gamma_3} \qquad (2.77)$$

$$= \gamma_1 \cdot \varphi_J(P_1) + \gamma_2 \cdot \varphi_J(P_1) + \gamma_3 \cdot \varphi_J(P_1)$$

(2.78)

WHERE EACH TERM IS AN INNER OR SCALAR PRODUCT AND $\Omega(P_1)$ IS QUANTIFIED. THE DIAGONALIZED VECTORS $\{\gamma_J\}$ ARE THE VECTORS OF OPTIMAL LENGTHS: P°_{11}, P°_{22}, P°_{33}. FOR OUR EXAMPLE THESE ARE ALL EQUAL AND OF MAGNITUDE 100. THESE OPTIMAL VECTORS ARE ORIENTED OR SHORTENED BY THE λ_{IJ} LAGRANGIAN PARAMETERS. ONE VIEWPOINT IS THESE ARE AS PROJECTIONS UPON THE P_3 VECTOR OR AXIS.

THEREFORE

$$\begin{Bmatrix} \Lambda_1 \\ \Lambda_2 \\ \Lambda_3 \end{Bmatrix} = \begin{Bmatrix} \lambda_{11} P_{11} \\ \lambda_{22} P_{22} \\ \lambda_{33} P_{33} \end{Bmatrix} = \begin{Bmatrix} (0.50).\ 100 \\ (0.60).\ 100 \\ (1.0\ ).\ 100 \end{Bmatrix} = \begin{Bmatrix} 50 \\ 60 \\ 100 \end{Bmatrix} \tag{2.79}$$

FOR PLOT II , TO FIND THE INTERCEPT VALUE FOR $\Omega(P_{2J})$

$$\frac{\partial \Omega(P_1)}{\partial P_2} \cdot \Big|_{P_{21}} = 0 = \frac{\partial}{\partial P_2} \left[\Lambda_1 \ \Delta_1 \right]_{P_{21}}$$

$$= \Lambda_1 \left(-a_{11} \cdot \frac{\partial P_1}{\partial P_2} - a_{21} \cdot 1 - a_{31} \frac{\partial P_3}{\partial P_2} \right) \Big|_{P_{21}^{\circ}}$$

$$+ \left(W_1 - a_{11} \cdot P_1 - a_{21} \cdot P_{21} - a_{31} \cdot P_3 \right) \cdot \frac{a_{21}}{a_{31}}$$

$$0 = 50 \left[-0.5 \ \left(\frac{-0.25}{0.50}\right) - 0.25 - 1.0\left(\frac{-0.5}{1.0}\right) \right]$$

$$+ \left[175 - 0.5\, P_1 - 0.25(60) - 1.0\, P_3 \right] \cdot (-0.25)$$

$$0 = 25 - 40 + 0.125 P_1 + 0.25 P_3$$

$$0 = -15 + 0.125 P_1 + 0.25 P_3$$

OR:

$$P_3 + 0.5\, P_1 - 60 = 0 \tag{2.80}$$

AS $P_1 \to 0$, $P_3 \to 60$ AS ITS INTERCEPT FOR J=2.

FOR PLOT I , TO FIND INTERCEPT VALUE FOR $\Omega(P_{1J})$ CURVE

$$\frac{\partial \Omega(P_1)}{\partial P_1} \Big|_{P_{11}^{\circ}} = 0 = \frac{\partial}{\partial P_1} \cdot (\Lambda_1 \Delta_1)_{P_{11}^{\circ}}$$

$$0 = \Lambda_1 \cdot \left(\frac{\partial \Delta_1}{\partial P_1}\right) \Big|_{P_{11}^{\circ}} + \left(\Delta_2 \cdot \frac{\partial \Lambda_1}{\partial P_1} \right) \Big|_{P_{11}^{\circ}}$$

$$0 = \left[\Delta_2 \left(P_{11}, P_2, P_3 \right) \right] \cdot (-0.5)$$

OR:

$$P_3 = -0.25\, P_2 + 125 \tag{2.81}$$

$= 125$ AS $P_2 \to 0$

VECTOR ARRAY

FIGURE 11-9

J → / i ↓	γ_1	γ_2	γ_3	$\Omega(P^{\circ}_{iJ})$
P°_{1J} FIXED	$(125-0.25P_2-1.0P_3)/\gamma_1 = \frac{\Delta_1(P^{\circ}_{11})}{\gamma_1}$	$(165-0.60P_2-P_3)/\gamma_2 = \frac{\Delta_2(P^{\circ}_{12})}{\gamma_2}$	$(255-1.1P_2-P_3)/\gamma_3 = \frac{\Delta_3(P^{\circ}_{13})}{\gamma_3}$	$(272.5-0.975P_2-1.5P_3).2 = \Omega(P^{\circ}_{1J})$
P°_{2J} FIXED	$(160-0.50P_1-P_3)/\gamma_1 = \frac{\Delta_1(P^{\circ}_{21})}{\gamma_1}$	$(200-P_1-P_3)/\gamma_2 = \frac{\Delta_2(P^{\circ}_{22})}{\gamma_2}$	$(255-2.0P_1-P_3)/\gamma_3 = \frac{\Delta_3(P^{\circ}_{23})}{\gamma_3}$	$(307.5-1.75P_1-1.5P_3).2 = \Omega(P^{\circ}_{2J})$
P°_{3J} FIXED	$(65-0.50P_1-0.25P_2)/\gamma_1 = \frac{\Delta_1(P^{\circ}_{31})}{\gamma_1}$	$(155-P_1-0.6P_2)/\gamma_2 = \frac{\Delta_2(P^{\circ}_{32})}{\gamma_2}$	$(310-2P_1-1.1P_2)/\gamma_3 = \frac{\Delta_3(P^{\circ}_{33})}{\gamma_3}$	$(265-1.75P_1-0.975P_2).2 = \Omega(P^{\circ}_{3J})$
$\sum = 2.\Omega(P_i)$	$2.(175-0.5P_1-0.25P_2-P_3) = 2\Delta_1(P^{\circ}_i) = K_1$	$2.(260-P_1-0.6P_2-P_3) = 2\Delta_2(P^{\circ}_i) = K_2$	$2.(410-2P_1-1.1P_2-P_3) = 2\Delta_3(P^{\circ}_i) = K_3$	$2.(845-3.5P_1-1.95P_2-3.0P_3) = \Omega(P_i).2$

WHERE $\Omega_0 = 845$ (2.82)

$$2\Omega(P_i) = \Omega(P^{\circ}_{1J}) + \Omega(P^{\circ}_{2J}) + \Omega(P^{\circ}_{3J}) \quad (2.83)$$

$$= K_1 + K_2 + K_3 \quad (2.84)$$

$$= 2.(845 - 3.5P_1 - 1.95P_2 - 3.0P_3) \quad (2.85)$$

FOR PLOT II

$$\Delta_1(P^{\circ}_{21}) + \Delta_2(P^{\circ}_{22}) + \Delta_3(P^{\circ}_{23}) = \Omega(P^{\circ}_{2J})$$

$$\gamma_1 \varphi_1(P^{\circ}_{21}) + \gamma_2 \varphi_2(P^{\circ}_{22}) + \gamma_3 \varphi_3(P^{\circ}_{23}) =$$

$$2.\,(307.5 - 1.75\,P_1 - 1.5\,P_3) = \qquad (2.86)$$

FOR PLOT I

$$\Delta_1(P^{\circ}_{11}) + \Delta_2(P^{\circ}_{12}) + \Delta_3(P^{\circ}_{13}) = \Omega(P^{\circ}_{1J})$$

$$\gamma_1 \cdot \varphi_1(P_{11}) + \gamma_2 \varphi_2(P_{12}) + \gamma_3 \varphi_3(P_{13}) =$$

$$2.\,(272.5 - 0.975\,P_2\ 1.5\,P_3) = \qquad (2.87)$$

REDUNDANT PLOT III

$$\Omega(P_1, P_2) = \Delta_1(P^{\circ}_{31}) + \Delta_2(P^{\circ}_{32}) + \Delta_3(P^{\circ}_{33}) = \Omega(P^{\circ}_{3J})$$

$$= \gamma_1 \varphi_1(P^{\circ}_{31}) + \gamma_2 \varphi_2(P^{\circ}_{32}) + \gamma_3 \varphi_{3(P^{\circ}_{33})}$$

$$= 2.(265 - 1.75\,P_1 - 0.975\,P_2) \qquad (2.88)$$

VECTOR ARRAY

FIGURE 11-10

i \ J	$\mathcal{V}_1$	$\mathcal{V}_2$	$\mathcal{V}_3$	$\Omega(P^{\circ}_{1J})$
P°_{1J} FIXED	$\varphi_1(P_{11})$	$\varphi_2(P_{12})$	$\varphi_3(P_{13})$	$\Omega(P^{\circ}_{1J})$
P°_{2J} FIXED	$\varphi_1(P_{21})$	$\varphi_2(P_{22})$	$\varphi_3(P_{23})$	$\Omega(P^{\circ}_{2J})$
P°_{3J} FIXED	$\varphi_1(P_{31})$	$\varphi_2(P_{32})$	$\varphi_3(P_{33})$	$\Omega(P^{\circ}_{3J})$
	$2.K_1$	$2.K_2$	$2.K_3$	$2.\Omega(P_I)$

WHERE $K_1 + K_2 + K_3 = 2.\ (\Omega_0 - B_1P_1 - B_2P_2 - B_3P_3)$ (2.89)

$$B_I = \sum_{J=1}^{3} \alpha_{IJ} \tag{2.90}$$

$$2.K_J = 2.\ \Delta_J(P_I) \tag{2.91}$$

$$2.K_1 = 2.\ (175 - 0.50P_1 - 0.25P_2 - P_3)$$
$$2.K_2 = 2.\ (260 - P_1 - 0.60P_2 - P_3)$$
$$2.K_3 = 2.\ (410 - 2P_1 - 1.1P_2 - P_3)$$

$$\dot{\Omega}_0 = \sum_{J=1}^{3} (\bar{W}_J) = 175 + 260 + 410 = 845 \tag{2.92}$$

$$\Omega(P_1) = \frac{K^T}{\delta} \cdot \underline{\delta} = \frac{K_1}{\delta 1} \cdot \delta_1 + \frac{K_2}{\delta_2} \cdot \delta_2 + \frac{K_3}{\delta_3} \cdot \delta_3 \tag{2.93}$$

AND $\underline{A}^T \cdot \underline{\delta} = \underline{B}$ AT EXTREMA. (2.94)

$$\underline{A}^T \cdot \underline{P}' = \frac{(\underline{K})}{\delta}' \quad \text{AT EXTREMA.} \tag{2.95}$$

STATISTICAL INTERPRETATION

SINCE $\delta_J = \left\{ (P^{\circ}_{11}) \cdot \lambda_{IJ} \right\}$ | FOR I=J = CONDITION WEIGHTED VALUE OF THE GENERAL POPULATION ARITHMETIC MEAN FOR P_I.

THEN $\delta_1 = (\bar{x}_1) \cdot \cdot (\frac{B_1}{B_3})$ (J=1) SET STATISTICAL REGRESSION COEFFICIENTS

$\delta_2 = (\bar{x}_2) \cdot (\frac{B_2}{B_3})$ (J=2) SET

$\delta_3 = (\bar{x}_3) \cdot (1)$ (J=3) SET AND GENERAL OR TOTAL POPULATION.

SINCE $W_J = (Y - B_0)_J = a_{1J} P_1 + a_{2J} P_2 + a_{3J} P_3$ (2.96)

$$\sum_{J=1}^{3} (\bar{W}_J) = B_1' \cdot \bar{X}_1 + B_2' \cdot \bar{X}_2 + B_3' \cdot \bar{X}_3 \quad (2.97)$$

$$= B_1 \left(\frac{B_3}{B_1} \cdot \hat{\gamma}_1\right) + B_2 \left(\frac{B_3}{B_2} \cdot \hat{\gamma}_2\right) + B_3 \hat{\gamma}_3$$

$$\frac{\sum_{J=1}^{3} (\bar{W}_J)}{B_3} = \hat{\gamma}_1 + \hat{\gamma}_2 + \hat{\gamma}_3 \quad (2.98)$$

FIGURES II -2 AND II-3 GRAPHICALLY PORTRAY THE VECTOR AND ALGEBRAIC RELATIONSHIPS OF (2.96), (2.97) AND (2.98). THE MAJOR ADVANTAGE OR VALUE OF THE MATHEMATICAL TECHNIQUE PRESENTED HERE IS SEEN TO BE ITS CLOSE TIE-IN WITH THE GRAPHICAL MODELS ,THEIR STATISTICAL INTERPRETATION AND EASY CONVERSION TO DESIGN CHARTS.

III THE ADAPTIVITY MODEL FOR THE WASTE- PROCESSING SYSTEM

THE TOTAL POPULATION DATA FOR 18 MONTHS WAS SEPARATED INTO SUBSETS OF SIX(6) -WEEK DURATION. EACH IS TREATED AS A J CONDITION SET. THE STATISTICAL OPTIMIZATION TECHNIQUES REPORTED IN OUR COMPANION PAPER --STUDY A -- WERE APPLIED TO EACH J SET. WE CONSIDERED A REDUCED PROBLEM OF SIX PARAMETERS - Y AND X_i FOR (i=1,2,3,4, AND 5). BOTH A LINEAR MULTIPLE PARAMETER REGRESSION ANALYSIS AND A PARABOLIC PLUS LINEAR AS A CURVILINEAR REGRESSION WERE MADE.

$$(Y)_J = (B_0)_J + \sum_{i=1}^{5} F_i^{(1)} (X_i) + \sum_{Q=6}^{10} F_Q^{(1)} (X_Q)$$

$$= (B_0)_J + \sum_{i=1}^{5} F_i^{(2)} (X_i) \quad (2.99)$$

SINCE $X_Q = (X_i)^2$ (2.100)

THE REDUCED SELECTION NUMBER OF VARIABLES WAS BASED UPON AN EARLIER ANALYSIS. THOSE VARIABLES EXHIBITED VERY LOW LINEAR CORRELATION COEFFICIENTS WERE TESTED AND REJECTED.

THE z_1^* CONFIDENCE TEST HAS BEEN APPLIED AND THE DATA TABULATED. THE % CONTRIBUTIONS OF EACH BASIC INDEPENDENT VARIABLE TO THE DEPENDENCY PARAMETER HAVE BEEN FOUND. TO ILLUSTRATE THE INFLUENCE OF THE ENVIRONMENTAL SETS OF CONDITIONS, A SERIES OF TRADE-OFF PLOTS FOR MULTIPLE $\underline{J}$ CONDITIONS HAVE BEEN DRAWN BY THE COMPUTER. A SECOND DEPENDENCY VARIABLE - AERATION RATE - WAS TREATED IN A SECOND STUDY OF THE SAME DIMENSIONS AND MAGNITUDE. WE WILL DISCUSS THE RESULTS BRIEFLY IN THIS PAPER AND REFER THOSE READERS INTERESTED IN SPECIFIC DETAILS TO THE ACCOMPANYING ENCLOSURE OF COMPUTER- PROCESSED, TABULATED AND DISPLAYED DATA.

Dependency Models of Sewage Flow upon Ambient Temperature

FROM THE REGRESSION MODEL ANALYSIS (BOTH LINEAR AND CURVILINEAR) WE HAVE REPLOTTED THE TRADE-OFF PORTRAITS FOR SEVERAL $\underline{J}$ SETS OF CONDITIONS. FIGURE II-11 AND FIGURE II-11' GIVE THE LINEAR CURVES FOR LOW AND HIGH $\underline{J}$ NUMBERS. FIGURE II-12 GIVES THE CONTRASTING CURVILINEAR. TO BETTER EXPLAIN THE SIGNIFICANCE OF THESE FAMILY PLOTS FIGURE II-14 IS INCLUDED. THIS BAR GRAPH GIVES THE ARITHMETIC MEAN VALUEFOR THE DAILY HIGH TEMPERATURE FOR EACH OF THE $\underline{J}$ TIME PERIODS AND COMPANION DATA SET.

IF WE WERE TO ASSUME THAT THE SUBSET DATA SAMPLE COULD BE CHARACTERIZED BY KNOWLEDGE OF THE SAMPLE ARITHMETIC MEAN, CONFIDENCE COULD BE ACQUIRED BY INSPECTION OF THE CURVILENEAR PLOTS. THE EXTREMUM FOR EACH CURVE OCCURS CLOSE TO THE COMPUTED SAMPLE MEAN . CERTAINLY , THE CURVILINEAR MODELS GIVE MORE REALISTIC

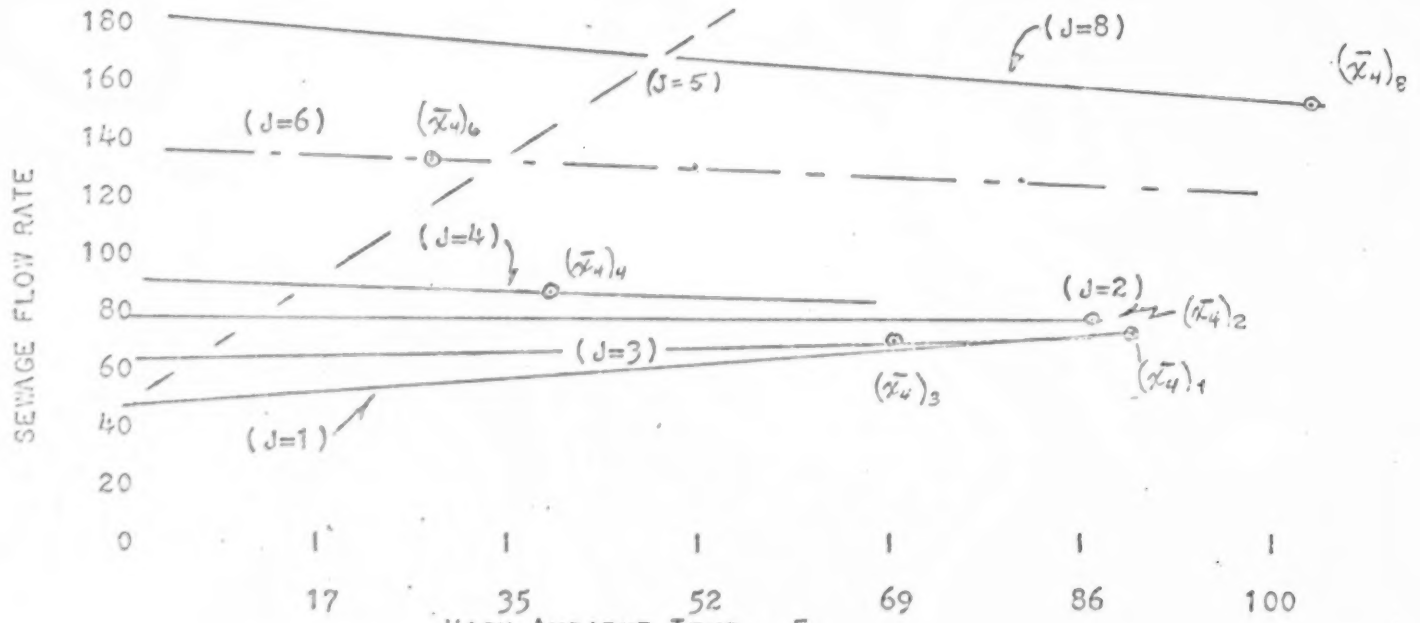


FIGURE II-11 LINEAR MODELS

REPRESENTATION TO THE SAMPLE SUBSETS. THE TOTAL POPULATION ARITHMETIC MEAN FOR EACH VARIABLE (X_i) REPRESENTS THE MATHEMATICAL OPTIMAL VALUE FOR THE ENTIRE SYSTEM.

TABLE II-1 AND TABLE II-2 ARE IN THE FORM OF SYSTEM DEFINITION MATRICES. THE INFLUENCE OR REGRESSION COEFFICIENTS ARE GIVEN AS THE CELL VALUES OF THE MATRIX FOR EACH J CONDITION ($J= 1,2,3,4,5,6,7,8$) AND FOR EACH OF THE PARAMETERS X_i ($i=1,2,3,4,5$) AND X_Q ($Q=6,7,8,9,10$). AGAIN, $X_Q=(X_i)^2$. THE DEPENDENCY VARIABLE FOR THE FIRST TABLE IS ($Y_{K=1}$) --THE SEWAGE FLOW RATE. FOR THE SECOND TABLE THIS VARIABLE IS ($Y_{K=2}$) --AERATION RATE. ONLY THE CURVILINEAR REGRESSION MODEL FOR EACH J SUBSET OF CONDITIONS AND DATA YIELDS REALISTIC TRADE-OFF PLOTS AND COMPANION REGRESSION MODELS. A REGRESSION MODEL AND ITS COMPANION ANALYSIS TOGETHER WITH CONFIDENCE CHECK ARE NEEDED FOR EACH J SUBSET. THE OPTIMAL PAY-OFF FUNCTION IS FOUND FROM UTILIZATION OF THE SET OF POPULATION ARITHMETIC MEANS ($\bar{X}_i$) FOR EACH VARIABLE AND THE SYSTEM MODEL COEFFICIENTS FOUND FROM THE COLUMN SUMS OF THE SYSTEM MATRIX (SEE 2.82 AND 2.85).

FIGURE II-13 SHOWS A CURVE INTERCONNECTING THE SUBSET (ARITHMETIC MEAN VALUES FOR X_4) WITH ALL OTHER X_i HELD AT THEIR POPULATION ARITHMETIC MEANS . THE ($\bar{X}_{RJ}$) DO OCCUR IN PROPER SEQUENCE (2-3-4-5-6) AND APPEAR TO FORM A REASONABLE CURVE--- PERHAPS IDENTIFIABLE AS TWO PARABOLAS WITH AN INFLECTION POINT NEAR (X_4) = 45 , THE POPULATION MEAN. THIS (X_R, X_{RJ}, X_I) IS A STATISTICAL FORM OF AN OPTIMAL SELECTION CURVE. IT CAN FORM THE BASIS OF EXTENSIONS TO DESIGN CHART FORMULATION.

FIGURE 11-12

CURVILINEAR

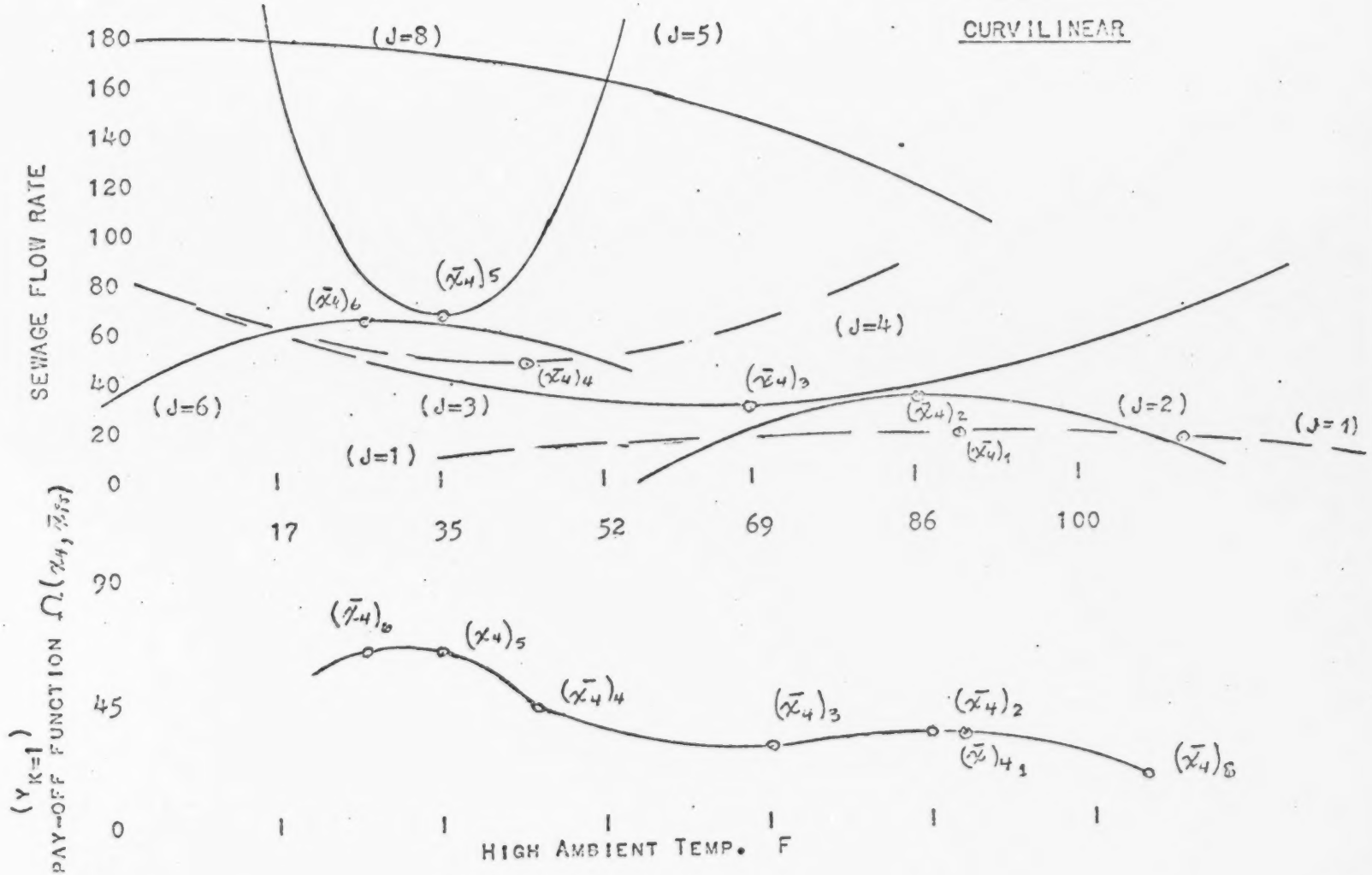


FIGURE 11-13

FIGURE II-14

Reproduced from best available copy.

J=	X_1	X_1^2	X_2	X_2^2	X_3	X_3^2	X_4	X_4^2	X_5	X_5^2	$Y_{K=1} - B_0$
1	0.0	-.704	-2080.	342.9	-420	49.3	580	-35.8	123.1	-.0092	-156238.29
2	0.0	.925	1660	270.4	-2262.4	22.2	104	-57.7	-5.77	.0003	178199.3
3	-3970.	0.0	-713.9	227.6	1916.6	-23.03	429.8	-1.86	-4.118	.0004	-174575.1
4	-429.5	1.15	-6681.	80.8	0.0	-2.21	-1452.8	14.9	-2.77	-.011	-175659.7
5	-8624	49.7	5727.	0.0	-31340.	194.6	-25757.	374.6	259.7	-.140	-1757733.5
6	9009	0.0	-73382	1019.	-1834.	0.0	5406.	-80.7	924.1	-.970	-210640.
7	0.0	463.5	154745	270.0	-11736.	119.5	-3064.	57.9	610.6	-.383	-657906.3
8	9756.	0.0	-63421.	773.3	5732.6	-66.4	-8236.7	70.6	816.4	-1.08	-292171.9
	5741.7	514.57	15854.1	2984.	-39943.2	293.96	-31094.7	341.94	2721.24	-2.5925	-3246725.49

Table II - 1

Sewage Flow = $(Y)_{K=1}$

97

	X_1 ($\times 10^{-4}$)	X_1^2 ($\times 10^{-9}$)	X_2	X_2^2	X_3	X_3^2	X_4	X_4^2	X_5	X_5^2 ($\times 10^{-4}$)	$Y_{K=2} - B_0$
J= 1	0.0	-1.0	4.019	-007	-4.718	.0515	-1.24428	.00636	-.0740	0.4	-80.479
2	0.4	-10.	4.01	-.0041	-.173	.00144	.4698	-.0026	.001	-.01	102.41
3	-1.0	20	2.77	.0135	-.122	.001	-.255	.0019	-.0007	.007	68.32
4	-10.4	300	.613	.024	-.621	0.0	1.47	-.022	.0397	-70	-47.31
5	-5.0	0.6	4.005	-.0122	.145	-.0016	-.129	+.0021	-.0018	1.8	46.39
6	-20.0	5	3.43	.0032	0.0	.0012	-.0528	.0008	-.018	0.2	54.62
7	-3.7	-.01	4.101	-.0158	.0763	-.0008	.0119	-.0003	-.0056	3.3	26.06
8	-5	1.6	4.123	-.021	-1.887	.0106	-.337	.00313	.0338	4.3	-59.32
	37.3	316.	27.07	-.0194	-7.2997	.06334	-.0661	-.0106	-.0256	-60.	110.691

Table II - 2

Aeration Rate = $(y_{K=2})$

IV Design Charts from Trade-off Studies

The transition from multiple statistical analysis and their companion trade-off studies under selected sets of conditions requires considerable effort and sound engineering judgment. Specifically, each potential interdependency for the selected set of subsystem and system parameters must be identified and quantified. The pay-off functions for selection of critical and significant design parameters must permit the addition of dimensions, practical tolerances and the expected resultant variations in system output variables due to these adopted tolerances.

The practical development of charts useful to the designer for selection of system and component sizes in all physical respects requires a down-to-earth understanding of system operation in total. Hence, part of our job as system analyst of an already existing and operating system is to assure that the final submitted conclusions are in the correct practical form. A trade-off study easily understood by a system analyst familiar with parameter dependencies may be so much confusion to a practical hardware designer. The transition across this explanation-understanding barrier presents a challenge greater than the analysis of this present most complex system. For example, the designer is interested in charts showing value ranges for :

1. Motor and rotor size, duty cycle, number of units versus aeration generation rates.
2. Limited ranges of transport flow rates for influent-affluent capacities.
3. Tolerances in design values required to permit adaptivity to expected rainfall for given

GEOGRAPHICAL COMMUNITY.

4. CAPACITY RELATIONSHIPS OF ALL VARIETIES VERSUS PHYSICAL PROPERTIES OF INFLUENT.
5. SETTLING TIME OF SYSTEM VERSUS DESIGN CAPACITIES TO HANDLE INFLUENT IMPULSES OF NUMEROUS VARIETIES (MECHANICAL, CHEMICAL, ENVIRONMENTAL)
6. EFFICIENCY FACTORS AND RANGES PROVIDED FOR PRACTICAL ADAPTIVITY TO EXPECTED RANGES OF EACH AND ALL ENVIRONMENTAL VARIABLES (HANDLED STATISTICALLY).
7. PRACTICAL RANGES OF SYSTEM EFFECTIVENESS FOR SEVERAL MODULAR SIZES OF PLANT.
8. PRACTICAL RANGES OF COST EFFECTIVENESS FOR SEVERAL MODULAR SIZES OF PLANT.
9. PREDICTION CHARTS FOR EXPECTED OPERATING COSTS, MAINTAINABILITY, COSTS FOR THE SEVERAL MODULAR SIZES OF PLANT.
10. CHARTS SHOWING EXPECTED PERFORMANCE IN ALL PRACTICAL ASPECTS FOR UTILIZATION THE SEVERAL MODULAR UNITS IN TANDEM, IN PARALLEL, AS ADD-ON UNITS TO OTHER OF PROCESSING SYSTEMS.

THE FURNISHING OF THIS LENGTHY LIST OF DESIRABLE GUIDES TO THE DESIGNER IS REALISTICALLY BEYOND THE SCOPE OF OUR PRESENT PROGRAM AND ASSIGNED STAFF. NEVERTHELESS, MOST OF THE INFORMATION AND AND INTELLIGENCE FOR SUCH TASKS MUST BE MADE AVAILABLE. HOPEFULLY, THESE OBJECTIVES WILL BE MEET IN THE THIRD YEAR OF THE PROGRAM.

ACKNOWLEDGEMENT

THE AUTHOR GRATEFULLY HAS RECEIVED VALUABLE ASSISTANCE FROM GRADUATE STUDENT-HARRY FAULKNER AND FROM THE STAFF OF THE OHIO UNIVERSITY COMPUTER CENTER IN CARRYING THE VERY EXTENSIVE SOFTWARE PROGRAMMING AND RESULTANT DATA PROCESSING TO PRACTICALITY.

REFERENCES:

A. Bartee, E. M. : " Statistical Methods in Engineering Experiments", Merrill Books, Cols.O. -1966.

B. Gamble, E. H. : " The Modern Engineering Method", Vol 1, part 2, Purdue University, School of Mechanical Engineering, - 1963.

C. Feller, William : " An Introduction to Probability Theory and its Applications", Vols. I and II. Wiley, 1966.

D. Gamble, E. H. : " A Critque on the Design of Complex Systems", copyrighted June 1967, Library of Congress Card No. A 927900

Study C

"Polynomial Regression Analysis of Input-Output Models"

(Applied to Domonstration
and Analysis of the
Somerset Oxidation-Ditch
Method for Waste-Water Processing)

by John Hungerford, BSME
E. H. Gamble, Ph.D.
Ohio University
Athens, Ohio

I. Background

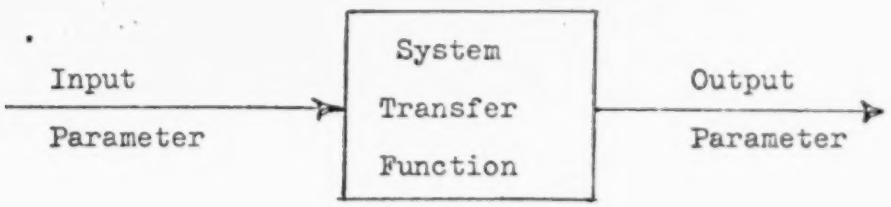
In our systems analysis study of the Oxidation-Ditch Method of waste-water processing, some decisions were required in selection of the limitations upon the complexity of the curvilinear multiple parameter regression anlaysis. The single - simple correlation coefficient gives some justification for assuming that those x_i having a $\rho_{x_i y}$ less than 0.1 have little influence upon the y variable. The magnitude of the total residual of errors tells the relative match of the set of $\{x_i\}$ for the regression model of y, but tells little about each x_i or the form of the dependence of y upon it, relative to the remainder set of x_i.

Methods such as the Reverse Markov Process and a(χ_i)* confidence test (reported in companion Study A - "Statistical Optimization") are sufficiently complicated and have such computer time needs that some more simple decision-process on regression complexity is needed. For this reason the polynomial regression analysis method for input - output complexity selection was developed and found to fill our specific needs.

One complete model study is provided as an example. A population of 100 observations was adopted from our

system measurement data. Polynomial regressions up to degree 6 were computed. Tabulated data results and a computer plot is furnished.

II The Method



The sewage treatment plant is viewed as an input-output model. A system parameter is sampled and measured at the input of the system. The parameter is again measured at the output of the system. The input parameter is selected as the independent variable and the output parameter is selected as the dependent variable. Polynomial regression techniques are used to develop the transfer function characteristics. The polynomial regression program was written to fit a 10th order polynomial. The algorithm increases the order of the polynomial by evaluating the improvement in the residual sum of squares. When no improvement is possible, the actual data is plotted on a graph. (These data points appear as 1's on the graph.) In addition, the theoretical curve of best fit is plotted. (These data points appear as 2's on the graph.)

Examples of the technique are provided. The results are summarized below:

Input Parameter	Output Parameter	Equation (Model)
Total PO_4 MG/L Raw Sample	Total PO_4 MG/L Final Sample	$Y=26.0012-1.9997X+.0921X^2-.0011X^3$
PH Raw Sample	PH Final Sample	$Y=-50.2404+14.6202X-.8243X^2-.0142X^3$
Solids Total MG/L-Raw Sample	Solids Total MG/L-Final Sample	$Y=-766.0579+4.1028X-.0037X^2$

III Discussion

The polynomial regression technique can be used to obtain relationships for the transfer function of the input-output model. This transfer function serves as a measure of the efficiency of the system for selected parameters. The appropriate software programming is

provided as part of the sample study. The major value of this short study has been to provide some practical justification for our limitation of the curvilinear regression analysis and models to parabolic form. The adaptivity analysis and resultant plots further strengthen this limitation choice.

REFERENCES:

A. Bartee, E. M. : " Statistical Methods in Engineering Experiments", Merrill Books, Cols.O. -1966.

B. Gamble, E. H. : " The Modern Engineering Method", Vol 1, part 2, Purdue University, School of Mechanical Engineering, - 1963.

C. Feller, William : " An Introduction to Probability Theory and its Applications", Vols. I and II. Wiley, 1966.

D. Gamble, E. H. : " A Critque on the Design of Complex Systems", copyrighted June 1967, Library of Congress Card No. A 927900